

**THE COST-EFFECTIVENESS OF
INTERVENTIONS TO SUPPORT SELF CARE**

**A THESIS SUBMITTED FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY**

BY

GERALD ANTHONY RICHARDSON

**DEPARTMENT OF HEALTH SCIENCES,
UNIVERSITY OF YORK**

DECEMBER 2007

CONTENTS

List of figures6

List of tables7

Abstract.....8

Acknowledgements.....9

Publications.....10

Conference presentations11

Seminar presentations.....11

Author’s Declaration11

Chapter 1. Economic evaluation of interventions to support self care..12

1.1 Background12

1.2 The decision problem and economic evaluation13

1.2.1 Who is the decision maker?13

1.2.2 What evidence exists at present for the decision maker?14

1.2.3 Which outcomes should inform the decision?14

1.2.4 What evidence should be used?14

1.3 Objectives.....15

1.3.1 Research questions15

1.4 Structure of thesis.....16

Chapter 2. Prevailing concepts in economic evaluation18

2.1 Introduction.....18

2.2 Economic evaluation in health care19

2.3 Normative foundations of economic evaluation19

2.3.1 Welfare economics, welfarism and economic evaluation20

2.3.2 Extra-welfarism and social decision making.....27

2.3.3 Concerns with the extra-welfarist approach.33

2.3.4 Welfarism, extra-welfarism and the economic evaluation of interventions to support self care34

2.3.5 Discussion of normative foundations of economic evaluation35

2.4 Single trial based economic evaluation and decision making36

2.4.1 Failure to consider all relevant alternatives37

2.4.2. Truncated time horizon38

2.4.3. Lack of relevance to decision context38

2.4.4. Failure to incorporate all relevant evidence.....39

2.4.5 Discussion of single trial based evaluations.....41

2.5 Statistical analysis and interpretation in economic evaluation42

2.5.1 Bayesians and Frequentists.....42

2.5.2 The role of inference43

2.5.3 Simulation in Bayesian cost-effectiveness analysis45

2.6 Discussion46

2.7 Conclusion.....47

Chapter 3. Systematic review of economic evaluations of self care support49

3.1 Introduction.....49

3.2 Evaluating self care support49

3.2.1 Comparator intervention.....49

3.2.2 Placebo effects.....50

3.2.3 Control group contamination50

3.2.4 Range of outcomes50

3.2.5 Patient costs.....51

3.2.6 Length of follow-up.....	51
3.2.7 Transferability of results	52
3.3 Methods.....	52
3.3.1 Self care support.....	52
3.3.2 Inclusion criteria	53
3.3.3 Search strategy.....	53
3.3.4 Checklist development.....	54
3.4 Results.....	54
3.4.1 Study characteristics	56
3.4.2 Study conclusions and quality	68
3.5 Discussion	70
3.6 Conclusion.....	71
Chapter 4. The Expert Patients Programme Cost-Effectiveness Evaluation.	73
4.1 Introduction.....	73
4.2 Background to the EPP	74
4.3 Methods.....	74
4.3.1 Participants	75
4.3.2 Intervention	76
4.3.3 Outcome measures.....	76
4.3.4 Perspective	78
4.3.5 Resource Use and imputation	78
4.4 Sources of data.....	79
4.4.1 Resource use.....	79
4.4.2 Unit costs	79
4.4.3 Health States and their value	80
4.5 Methods of analysis	82
4.5.1 Missing data and imputation.....	82
4.5.2 Incremental cost-effectiveness ratios (ICERs) and net monetary benefits	83
4.5.3 Time horizon	84
4.5.4 Sensitivity analyses.....	85
4.6 Results.....	85
4.6.1 Missing data	85
4.6.2 Resource use	86
4.6.3 Unit costs	90
4.6.4 Health States.....	91
4.6.5 Health state and self-efficacy	92
4.6.6 Quality-adjusted life-years (QALYs) based on imputed data.....	93
4.6.7 Total cost	94
4.6.8 Incremental Cost-Effectiveness Ratio (ICER)	96
4.7 Net Monetary Benefits and CEAC	96
4.8 Sensitivity analysis.....	97
4.8.1 Higher intervention costs.....	97
4.8.2 Excluding patient costs	97
4.8.3 Maintenance of improvements	98
4.8.4 Multiple imputation	98
4.8.5 Adjustment for condition, age and gender.....	99
4.8.6 Conventional (classical) statistical analysis.....	100
4.9 Discussion	100

4.10 Conclusion	102
Chapter 5. Evidence synthesis of trials with Individual Patient Data (IPD) of chronic disease self management programme (CDSMP)	105
5.1 Introduction	105
5.2 Trial characteristics	106
5.2.1 Trial results	107
5.3 IPD meta-analysis	107
5.4 Methods	109
5.4.1 Trial characteristics	110
5.4.2 Meta-analysis of IPD	111
5.4.3 Modelling assumptions	113
5.4.4 Sensitivity analyses	119
5.5 Results	119
5.5.1 Choice of covariates for the model	119
5.5.2 Distributional assumptions	121
5.5.3 Cost-effectiveness	122
5.5.4 Sensitivity analysis	123
5.5.5 Checks for convergence and auto-correlation	124
5.6 Discussion	125
5.7 Conclusion	126
Chapter 6. Evidence synthesis of IPD and summary statistics of surrogate outcomes	127
6.1 Introduction	127
6.2 Modelling concepts	129
6.2.1 Synthesising IPD and AD	129
6.2.2 The use of surrogate outcomes and the assumption of conditional independence	130
6.2.3 Exchangeability	132
6.2.4 Fixed and random effect models	134
6.2.5 Uncertainty in economic evaluation	134
6.3 Model development	135
6.3.1 The “direct” model	137
6.3.2 The “indirect” model	138
6.3.3 The “mixed” model	139
6.3.4 Model estimates	140
6.3.5 Summary of the models	141
6.4 Methods	141
6.4.1 Model assumption	141
6.4.2 Search methods	142
6.4.3 Software	142
6.5 Results of search for relevant clinical trial literature	144
6.5.1 Inclusion and exclusion of studies	144
6.5.2 Review of included studies	145
6.5.3 Clinical trial summary	148
6.6 Model Results	149
6.7 Discussion	153
6.8 Conclusion	158
Chapter 7. What outcomes do patients value? A discrete choice experiment	160
7.1 Introduction	160

7.1.1 Background	160
7.1.2 Aims and objectives	161
7.2 DCEs in the health care evaluation literature.....	162
7.3 Design of DCE.....	164
7.3.1 Full factorial and fractional factorial designs	164
7.3.2 Properties of DCE design.....	165
7.3.3 Interaction terms and length of questionnaire	166
7.4 Theoretical basis for DCEs	167
7.5 Econometric estimation techniques for DCE	169
7.6 Methods.....	171
7.6.1 Defining attributes and their levels	171
7.6.2 Questionnaire methods	172
7.6.3 Consistency.....	174
7.6.4 Pilot study.....	176
7.6.5 Main study.....	176
7.6.6 Main study sample	176
7.6.7 Model estimation	177
7.6.8 Incorporating the estimates of the value of self-efficacy into the CEAC	177
7.6.9 Sensitivity analyses.....	178
7.7 Results.....	178
7.7.1 Interaction terms	180
7.7.2 Standard probit model with constant term.....	180
7.7.3 Random effects probit.....	182
7.7.4 Generalised Linear and Latent Mixed Models (GLLAMM)	183
7.7.5 Random effects clustered on question number.....	185
7.7.6 Sensitivity analyses.....	186
7.7.7 Impact of inclusion of self-efficacy in the CEAC	187
7.8 Discussion	189
7.9 Conclusions	192
Chapter 8. Discussion.....	193
8.1 Summary of individual chapters.....	193
8.2 What could have been done differently?.....	200
8.3 Recommendations for future research.....	203
8.4 Conclusion.....	204
References.....	205
Appendices.....	220

List of figures

Figure 3.1 Selection of papers for assessment against criteria list.....55

Figure 4.1 The EQ-5D instrument81

Figure 4.2 Calculation of a QALY82

Figure 4.3 CONSORT diagram of the EPP national evaluation86

Figure 4.4 QALYs vs self-efficacy93

Figure 4.6 Scatter plot of incremental costs and effects of EPP.....104

Figure 5.1. Histogram of untransformed total costs of EPP and Griffiths combined
.....116

Figure 5.2. Histogram of (natural) log costs116

Figure 5.3 Histogram of distribution of QALYs form the 2 datasets combined ...117

Figure 5.4 Cost-effectiveness acceptability curves with two datasets with
alternative distributional assumptions123

Figure 6.1 Potential mechanisms for CDSMP to impact on costs/QALYs136

Figure 6.2 The “direct” model138

Figure 6.3 The “indirect” model139

Figure 6.4 CEACs for models using alternative mechanisms.....151

Figure 6.6 costs and effects using the direct model (above)155

Figure 6.7 costs and effects using the indirect model155

Figure 7.1 Example question from discrete choice experiment.....175

Figure 7.2. Impact of including self-efficacy on probability of EPP being cost-
effective.....189

List of tables

Table 3.1. Summary of economic evaluation papers	57
Table 3.2 Summary performance of economic evaluation papers against checklist	65
Table 4.1 Baseline demographic and health characteristics	77
Table 4.2: Mean resource use in the two groups over the 6-month period	87
Table 4.3: Mean resource use in the two groups over the 6-month period based on complete case analysis	89
Table 4.4: Unit costs of resources used	90
Table 4.5. % of patients in each EQ-5D dimension by group at baseline and six month follow-up.....	91
Table 4.6: Mean EQ-5D score at baseline and follow-up by group	92
Table 4.7 Unadjusted and adjusted mean change in QALYs per patient.....	94
Table 4.8 Total costs per patient in the two groups.....	94
Table 4.9. Major cost items by group	96
Table 4.10 Sensitivity analyses	99
Table 5.1 Comparison of baseline characteristics of two IPD trials.....	106
Table 5.2 Incremental costs and QALYs over 12 months in the two trials	107
Table 5.3 Regression of potential explanatory variables on QALYs.....	119
Table 5.4 Stepwise elimination model for QALYs	120
Table 5.5 Comparison of alternative distributional assumptions on model fit.....	121
Table 5.6 Comparison of models including covariates	122
Table 6.1 Direct, indirect and mixed models	143
Table 6.2 Summary of studies included in the analysis.....	149
Table 6.3 Results from three models with 95% confidence intervals where appropriate.....	149
Figure 6.5 EVPI of alternative models.....	154
Table 7.1 Attributes and levels used in the DCE	172
Table 7.2. Characteristics of non-responders and responders type.....	179
Table 7.3. Interaction terms in the DCE	180
Table 7.4. Probit model with constant term	182
Table 7.5. Random effect probit model with constant term	183
Table 7.6 GLLAMM model	184
Table 7.7 Random effects on question number.....	185
Table 7.8 Responses to short and long questionnaires	187
Table 7.9. Inclusion of “inconsistent” responses	187
Table 7.10. Numbers moving between self-efficacy category by group.	188

Abstract

There is considerable policy interest surrounding the introduction of interventions to support self care in the UK. Advocates of these interventions claim that patient outcomes are improved and costs to the health service could be reduced.

This thesis critically reviews the published evidence around the cost-effectiveness of interventions to support self care, and uses the example of the Expert Patients programme (EPP) to analyse the cost-effectiveness of these interventions. An economic evaluation conducted alongside a Randomised Controlled Trial (RCT) of the EPP is considered as the starting point for the analysis. Further potentially relevant evidence is then incorporated including other UK based randomised trial evidence, non UK evidence and non-randomised evidence. The assumptions required for each of these models, as well as the impact of model selection on the adoption decision and the value of future research is considered in an explicit framework.

The analyses above are conducted in the extra-welfarist tradition where “health” is considered to be the maximand in the decision makers’ objective function. The inclusion of other outcomes that may not be incorporated in “health”, such as levels of isolation and process measures, such as speed of access to health professionals, is also considered. How these additional outcomes could be incorporated into a cost-effectiveness framework is explored.

The strengths and weaknesses of the analyses are discussed and recommendations for future research are presented.

Acknowledgements

The completion of this thesis was made possible by a fellowship from the Department of Health National Co-ordinating Centre for Research Capacity Development. I would like to thank them for this assistance.

In addition, I am very grateful to my supervisor Professor Mark Sculpher, who has contributed useful comments and advice throughout the duration of this project. Members of the Thesis Advisory Group, Professor Trevor Sheldon and Professor Ian Watt have also had useful input, for which I am thankful.

Members of the academic community within the Centre for Health Economics are also acknowledged. In particular, Helen Weatherly and Hugh Gravelle contributed to an earlier draft of Chapter 3, Neil Hawkins contributed to the analysis in Chapters 5 and 6 and Chris Bojke contributed to Chapter 7. In addition, conversations and advice from Karl Claxton, Stefano Conti, Susan Griffin, Laura Bojke, Steve Palmer, Nigel Rice and Andrea Manca have been invaluable. Other individuals from other academic centres, in particular Dr Karen Bloor, Prof Christine Godfrey (from the Department of Health Sciences, University of York) Pete Bower, Anne Rogers, David Reeves and Anne Kennedy (from the National Primary Care Centre for Research and Development (NPCRDC) at the University of Manchester), Alex Sutton from the University Of Leicester, Chris Griffiths of Centre for Health Sciences, Barts and The London, and Katherine Payne from the University of Manchester have also supported this work. Comments for colleagues at Health Economists Study Group (HESG) have also been useful.

I would also like to acknowledge Frances Sharp for her painstaking proof reading and Liz Ingham for help with preparing this document at various stages.

I would also like to take this opportunity to thank three referees, Professor Miranda Mugford (University of East Anglia), Professor Brendan Delaney (University of Birmingham) and Professor David Torgerson (University of York) for their very useful discussion and comments.

During the period of this fellowship, I spent some time abroad and would like to thank in particular Professor Wendy Chaboyer of Griffith University, Queensland, Australia for her support (both academic and practical) during this period.

Finally, I would like to acknowledge the help, advice and continual support of Su Golder who encouraged me throughout the duration of this project.

This thesis is dedicated to my Mum and Dad, who I know will be proud of this achievement.

Publications

I have made efforts to disseminate the findings of the research presented in this thesis as they were undertaken. The following papers, based on work in this thesis, have been published:

Kennedy A, Reeves D, Bower P, Lee V, Middleton E, Richardson G, Gardner C, Gately C, Rogers A. The effectiveness and cost effectiveness of a national lay led self care support programme for patients with long-term conditions: a pragmatic randomised controlled trial. *Journal of Epidemiology and Community Health* 2007;61:254-261

Richardson G, Kennedy A, Reeves D, Bower P, Lee V, Middleton E, Gardner C, Gately C, and Rogers A. Cost-Effectiveness of the Expert Patient Programme (EPP) for patients with chronic conditions. *Journal of Epidemiology and Community Health* (in press). 2007. 10.1136/jech.2006.057430. accepted Jun 2007

Richardson G, Gravelle H, Weatherly H and Ritchie G. Cost-effectiveness of interventions to support self-care- a systematic review. *International Journal of Technology Assessment in Health Care* 2005; 21(4) 423-432.

The following papers are in submission:

Richardson G, Bojke C, Kennedy A, Reeves D, Bower P, Lee V, et al. What outcomes are important to patients with long term conditions? A discrete choice experiment. *Value in Health* 2007.

Conference presentations

Health Technology Assessment International. Presented paper “Are Interventions promoting Self-management cost-effective?”. May 31-June 3rd 2004. Krakow, Poland.

The Cochrane Colloquium. Presented paper “Using a Discrete Choice Experiment (DCE) to assess patients’ valuation of psychological outcomes”. October 21-26th 2005, Melbourne, Australia

Health Economists Study Group. Presented paper “Incorporating “psychological” outcomes into cost-effectiveness analysis”, York 2006

The Cochrane Colloquium. Presented paper “What outcomes are important to patients with long term conditions?”. October 22-27th 2006, Dublin, Ireland

Health Economists Study Group. Presented paper “A comparison of direct, indirect and joint analyses of cost-effectiveness.” Brunel January 2007

Seminar presentations

Valuation of psychological outcomes using discrete choice experiments. Presentation at Economic Evaluation workshop. University of York. June 2005

Do patients value psychological outcomes and should they be incorporated into cost-effectiveness analysis? Presentation at University of Manchester, June 2006

Q methodology and Discrete Choice in Preference Elicitation. Department of Health sciences, University of York, November 2006.

Do patients value psychological outcomes and should they be incorporated into cost-effectiveness analysis? Presentation at University of Brunel, January 2007

Synthesis of Individual Patient Data (IPD) and Aggregate Data (AD). Presentation at Economic Evaluation workshop. University of York. June 2007

Are interventions to support self care cost effective? Presentation to the Alliance for Self Care. University of Dundee, December 2007

Author’s Declaration

The views expressed in this thesis are those of the author and not necessarily those of the funders. The author declares that the research in this thesis is the result of their own investigation, other than those acknowledged previously, and that none of the material contained in this thesis has previously been submitted for a degree in this, or any other, awarding institution.

Gerald (Gerry) Richardson 2007

Chapter 1. Economic evaluation of interventions to support self care

1.1 Background

Chronic conditions account for around two-thirds of the global burden of disease,¹ with 32% of the adult population of Great Britain suffering from a long standing condition.² There is a consequent interest by policy makers to address the management of these conditions. In the UK, the National Health Service (NHS) policy for patients with chronic conditions who are at “low risk” has been to encourage these individuals to “self-care”,³ and this is considered central to the management of these conditions.⁴ The Wanless report described self care as one of the principal determinants of efficiency of the NHS.⁵ As a consequence, interest in the effectiveness of interventions to support individuals in their self care has increased in recent years.⁶ The cost-effectiveness of these interventions has attracted less interest, and evidence in favour of the increased use of interventions to support self care has been based on poorly conducted studies generating unreliable conclusions.⁷

Self care has been defined as ‘the care taken by individuals towards their own health and well being: it comprises the actions they take to lead a healthy lifestyle; to meet their social, emotional and psychological needs; to care for their long-term condition; and to prevent further illness or accidents’.³ A range of interventions are available to support individuals to self care in a variety of settings and for a range of conditions or ailments.³

The NHS Plan identified self care as an important factor in providing a “patient centred health service”.⁸ More recently a Department of Health report has claimed that supporting self care can “improve health outcomes [and], improve patient satisfaction”,³ while the roll-out of self care support programmes “will improve the length and quality of lives”⁹ and “could create a generation of patients empowered to take action to improve their health”.¹⁰

There are many interventions available to support or enhance self care. There is differing quality and quantity of evidence regarding their effectiveness. However, in a budget constrained system such as the National Health Service (NHS), it is important for these interventions to promote patients' self care or self management to be *cost-effective* as well as clinically effective; otherwise scarce resources will be used on these interventions that could be better spent elsewhere in the system. The use of economic evaluation as a discipline to tackle this constrained optimisation problem are described in the next two chapters. The methodological challenges associated with the use of economic evaluation depend upon the context in which the decision is to be addressed. For example, "Who decides what interventions are cost-effective, on behalf of whom, using which evidence?" These issues are discussed in more detail in the next section.

1.2 The decision problem and economic evaluation

The decision problem addressed in this thesis is to assess whether these interventions provide additional benefits that are worth paying for, relative to other appropriate comparators, in the budget constrained system of the UK NHS. Economic evaluation is an explicit methodology for addressing this decision problem. In order to establish the relative cost-effectiveness of interventions to support self care, it is necessary to explicitly estimate the incremental costs and effects of providing these interventions compared with suitable alternatives.

1.2.1 Who is the decision maker?

The decision problem described above appears uncontroversial. However, it raises the question of who is the decision maker and where he/she receives his/her legitimacy from? Traditional welfarist economists would not require a decision maker as the sum of individual utilities is sufficient to determine the "best" option. The use of welfarist economics in the health sector has been questioned and the social decision making approach using extra-welfarist (or non-welfarist) concepts have become standard practice in economic evaluation. Simply put, this approach considers that the political process delivers an entity that can either act as a legitimate decision maker itself or delegate power to another body which then has a legitimate role.

1.2.2 What evidence exists at present for the decision maker?

It is feasible that there is sufficient evidence to make a decision on the cost-effectiveness of interventions to support self care in the literature published to date. The decision problem stated above may be solved by examination of the existing literature if that literature consisted of reliable and relevant evidence using appropriate measures of outcome.

1.2.3 Which outcomes should inform the decision?

Economic evaluations consider incremental costs and benefits of an intervention(s) compared to a suitable alternative. However, to be most useful for a decision maker facing the decision problem above, outcomes need to be comparable across conditions. This requires the use of a generic measure of health related quality of life, such as the quality-adjusted life-year (QALY). Where the analysis uses a non-generic (or disease specific) outcome measure, the decision maker may be faced with a range of outcomes that he/she is not familiar with or cannot value easily. While the intervention may improve patient outcomes as estimated by these clinical measures, if the intervention is also cost-additive, decision making becomes problematic. The limitation of generic instruments is that all the outcomes of interest may not be picked up by the instrument, either because it is not a health improvement or because the instrument is not designed (or sensitive enough) to detect certain health dimensions.

1.2.4 What evidence should be used?

A recurring theme of this thesis is the inclusion of “relevant” evidence and what constitutes relevance. The decision problem described above is focussed on the UK NHS. However, it is possible (or likely) that evidence from outside the UK could be of interest and should be used to inform the analysis and hence the decision problem.

As well as different geographical locations, studies may be of varying quality, using different outcome measures, in different patient populations with a range of conditions and severity of conditions. Depending on the definition of relevance, all or none of these studies could be included.

1.3 Objectives

The discussion above invites a series of objectives and research questions which are examined in the remainder of this thesis. These objectives and research questions are as follows:

- To identify the appropriate paradigms and forms of analysis for a cost-effectiveness study of a specific intervention to support self care
- To identify and critically appraise existing economic evaluations of interventions to support self care and identify appropriate methodology to inform a cost-effectiveness study of an intervention to support self care.
- To assess the cost-effectiveness analysis of one well-documented intervention designed to support self care.
- To incorporate evidence from another cost-effectiveness study based on a randomised controlled trial (RCT) conducted in the UK.
- To expand the evidence base on interventions to support self care to include evidence from non RCT data outside the UK.
- To expand the measure of outcome beyond health
- Examine the possibility of including expanded outcomes in cost-effectiveness analysis.

1.3.1 Research questions

- What are the appropriate paradigms for the design conduct and analysis of economic evaluation?
- Does the existing evidence base provide sufficient evidence to answer the specified decision problem?
- What lessons can we learn from the existing literature to help in the design, conduct and analysis of economic evaluations of interventions to support self care?
- Does the incorporation of evidence from other sources impact on the results, conclusions and recommendations for future research?
- Should health always be the outcome of interest and how does the expansion of the outcome measure impact on the decision problem?

1.4 Structure of thesis

Chapter 2 examines the schools of thought within economic evaluation. The distinction between welfarist and extra-welfarist position and the influence that the selection of paradigm may have on the design, conduct and analysis of economic evaluation is examined. Other controversies with the economic evaluation discipline including the statistical paradigm of choice, and the use of evidence synthesis in economic evaluation are also discussed.

In Chapter 3 the existing literature on economic evaluation of self care support interventions is identified, reviewed and appraised. The quantity and quality of this evidence is used to assess whether there are any studies that may inform the decision problem presented above. In addition, this literature may provide lessons or recommendations for the future use of economic evaluation in self care support interventions.

Chapter 4 uses the evidence of Chapters 2 and 3 to inform the design, conduct and analysis of a single trial based evaluation of an intervention to support self care, the Expert Patients Programme (EPP). The limitations of this single trial based approach, in terms of ignoring potentially relevant information are then discussed.

Chapter 5 then extends the analysis of Chapter 4 by incorporating individual patient data from another UK based randomised controlled trial (RCT) of the intervention used in the EPP, while Chapter 6 extends the evidence base further by incorporating evidence from non-UK studies, including non RCT evidence. The impact these additional data have on results, conclusions and recommendations are discussed.

Chapters 4, 5 and 6 use a generic measure of outcome, the QALY, enabling decision makers to make judgements that are consistent across analyses. However, there is concern in some quarters that the QALY may not pick up all outcomes of interest. For the evaluation of interventions to support self care, other outcomes have been postulated as being important. The relative values of these

outcomes, as compared with health related quality of life, is presented as a Discrete Choice Experiment (DCE) in Chapter 7.

The final chapter, Chapter 8, presents an overview and discussion of the previous chapters. The principal findings of each of the chapters are summarised, together with their respective strengths and weaknesses. The implications of each chapter for the decision problem are discussed and issues arising from this thesis that may form an agenda for future research in this area are also be presented.

Chapter 2. Prevailing concepts in economic evaluation

2.1 Introduction

Chapter 1 outlined the recent policy debate and the decision problem facing decision makers. Economic evaluation provides a methodology that informs the decision problem by establishing whether interventions to support self care are cost-effective (or which of these interventions are cost-effective) and should therefore be considered as providing value for money in a budget constrained system. However, economic evaluation is an evolving discipline and is not free from controversy. This chapter aims to describe ongoing controversies and competing paradigms within economic evaluation and their importance in addressing the decision problem. Inevitably, analysts must make choices between paradigms. The choice of paradigms is important. The conduct of the evaluation and analysis of data are dependent on the paradigm selected. This influences the results and conclusions of any analysis. Ultimately, the recommendations for both the decision to adopt a strategy/intervention and the related decision of whether to conduct more research can depend on the choice of paradigm. This chapter also seeks to justify the choices made between the various paradigms that will be used in the remainder of this thesis.

This chapter commences with a brief summary of the requirement for economic evaluations of self care support interventions. The normative foundations on which economic evaluation are based is then discussed (including welfarism and extra-welfarism) together with limitations of each approach. The historical use of single trial based evaluations is described, and the potential limitations of this approach are considered.

Finally, the analysis and interpretation of results of economic evaluation is considered in both Bayesian and frequentist paradigms and the use of statistical inference as a decision making tool is debated.

2.2 Economic evaluation in health care

Health service decision makers in the UK and elsewhere are increasingly requiring evidence of cost-effectiveness before making the funding decision.¹¹⁻¹⁴ While their use has been primarily in the assessment of the cost-effectiveness of new pharmaceuticals,^{15, 16} economic evaluation has an important role in the assessment of cost-effectiveness of health promotion programmes and interventions.¹²

The burden of disease associated with chronic conditions was described in Chapter 1. There are potentially substantial implications (both costs and benefits) associated with the roll-out of large programmes to treat these conditions, such as the Expert Patients Programme (EPP), across the NHS. Given the size of the target population and the potential benefits and costs of rolling out such programmes, it is surprising that this intervention has not been subjected to rigorous evaluation previously. Nevertheless, this lack of rigorous evidence means that there is considerable potential for economic evaluation to be conducted in this sphere to assess both the cost-effectiveness of these interventions and the value of conducting further research in this area.

However, economic evaluation requires a theoretical framework. Without an appropriate framework, it is not possible to deliver a consistent and useful product for decision makers. There is considerable debate around the appropriate framework for economic evaluation, and these are discussed below.

2.3 Normative foundations of economic evaluation

Though there are a number of theoretical “schools of thought” amongst health economists working in economic evaluation, the debate between welfarists and extra-welfarists is perhaps the most contentious. Though there is not complete consensus as to the contents of either school¹⁷ the principal characteristics of the two paradigms are outlined briefly below.

2.3.1 Welfare economics, welfarism and economic evaluation

Positive economics, which explains causes and effects in economic variables (such as price and demand) makes no value judgements, but is silent on recommending which policy/decision *should* be made. In contrast, welfare economics is normative in nature in that value judgements are incorporated and therefore allows the comparison of various states of the economy (or part thereof) and what it should look like. Welfare (or welfarist) economics is based on four central beliefs,^{18, 19} namely that individuals maximise their utility, that individuals are the best judges of their own utility (individual or consumer sovereignty), that utility is a consequence of behaviour and that utility information *only* is relevant in making decisions about what is best. This final tenet is frequently termed “welfarism”. Sen has called this the “evaluative space”,²⁰ and the assertion is that only utility of individuals can contribute to welfare. This is where much of the confusion around definitions occurs. Welfare economics incorporates all the above four tenets, of which welfarism is one of the four. Later authors such as Culyer (1991) reduced the four tenets to two principles :²¹

- a) social welfare is a function of individual utilities
- b) individual utilities are functions of the goods and services consumed by those individuals.

These tenets undoubtedly generate a coherent theory that can address the issue of whether changes (either within the health sector or between the health sector and other sectors) improve social welfare. In its strictest interpretation, interpersonal comparisons of utility are not required or permitted.

However, strict interpretation of this theory allows that there is only an improvement in social welfare if at least one person is made better off without any other being made worse off (a Pareto improvement). This is a restriction that would prevent any social decision making, as undoubtedly some are made worse off with any such decision. For example, the provision of improved health care provision for all individuals implies that either other sectors (such as education) would experience budgetary cuts (which would make some worse off) or there

would have to be an increase in taxation (or alternative mechanism) which would make tax-payers worse off.

This limitation of the Pareto improvement led to the separation of the efficiency (the size of the cake) and distributional aspects (how the cake is divided) of social welfare (see for example Sugden and Williams).²² If the distributional aspects are not considered, or left to political processes, a *potential*/Pareto improvement exists where the “winners” from a change can in theory compensate the “losers” and remain better off than before. In practice, the Kaldor-Hicks compensation test is often used as the test for potential Pareto improvements and is the basis for cost-benefit analysis. It should be noted that all Pareto improvements would pass the Kaldor Hicks compensation test but that the reverse does not hold. Few examples of Kaldor Hicks improvements would be Pareto improvements (as someone is likely to be worse off). A potential problem of reversibility of this criterion (known as the Scitovsky paradox) is discussed in section 2.3.1.1 below.

Allowing potential Pareto improvements permits a much broader range of policy options/decisions to be considered. However, the downside of this enhanced ability to inform decisions is that issues of equity²³ (and some would argue morality,²⁴) are not incorporated.

2.3.1.1 What is in the decision maker's objective function under welfarism?

Culyer²¹ has suggested that the assumptions underpinning the welfarist approach are restrictive, and that social welfare is a function of more than individual utilities (for example social welfare is likely to include concepts of justice and fairness) and that these individual utilities are not based solely on goods and services consumed but also on other factors such as process utility.²⁵ Several commentators, see for example Mooney²⁶, question the use of the individual as the appropriate unit of analysis. Communitarians, for example, consider the community and focus on social cohesion and well-being rather than the rights or utilities of individuals. In this approach the “community” decides what constitutes a “good” and makes decisions to prioritise between goods.

Sen, in his capabilities approach,²⁷ considers human well-being to be based on both activities and capabilities (the opportunities an individual has to achieve these activities, rather than purely utility). Thus the capabilities approach includes the capacity to exercise freedom of choice and the capacity to live to old age among other non-utility concepts. He considers human welfare to be multi-dimensional rather than based on a single concept of utility. The deprivation of these capabilities through oppression, poverty, ill health or other factors is considered to reduce “well-being”.

Sen²⁸ is also concerned with the impact of welfarism on equity and states that “concern with equity must militate against the use of utilitarianism”. The desire for equity can take many forms. Sen’s paper, *The Impossibility of a Paretian Liberal*,²⁹ asserts that even a minimal degree of liberalism is inconsistent with the use of Pareto optimality as a mechanism for social decision making. This example is a more specific version of Arrow’s Impossibility theorem. Arrow demonstrated that aggregating preferences of more than two members of society with a three choice decision cannot be achieved without the existence of unfairness or inconsistency.³⁰ More specifically aggregation of preferences with three decisions is not possible without contradicting one of the following:

- ☐ monotonicity (more of a good is preferred to less)
- ☐ independence of irrelevant alternatives (if A is preferred to B, the introduction of C should not make B preferable to A)
- ☐ non-dictatorship (the welfare function represents more than the welfare function of one individual)
- ☐ non-imposition (societal preferences can be ranked from the set of individual preferences)
- ☐ unrestricted domain (all preferences of all members of society should yield a complete ordering for all choices).

Sen summarises these arguments.³¹ He presents a classic example, showing that even when utility information is complete, welfarism and other utility based criteria could lead to a situation where torture is preferred to no torture (see example

below). Clearly, torture can be substituted for any other “undesirable” characteristic of a society.

“Consider a set of three social states x, y and z, with the following utility numbers for persons 1 and 2 (there are no others).

	x	y	z
Person 1's utility	4	7	7
Person 2's utility	10	8	8

“In x person 1 is hungry while 2 is eating a great deal. In y person 2 has been made to surrender a part of his food supply to 1. While 2 is made worse off, 1 gets more utility, and the sum total of utility happens to be larger (with diminishing marginal utility).

“It is clear that y must be judged to be better than x by utilitarianism (since the utility sum is larger for y), by the so called 'Rawlsian maximin' or its lexicographic extension 'leximin' (since the worst-off person's utility is larger in y than in x), and indeed by virtually all the equity criteria that have been proposed in the literature using utility data. Let us take y to be better than x.

“Consider now z. Here person 1 is still just as hungry as in x, and person 2 is also eating just as much. However, person 1, who is a sadist, is now permitted to torture 2, who - alas - is not a masochist. So 2 does suffer, but resilient as he is, his suffering is less than the utility gain of the wild-eyed 1. The utility numbers in z being exactly the same as in y, welfarism requires that if y is preferred to x, then so must be z. But y is socially preferred to x. So z is preferred to x as well, thanks to welfarism.

“The conclusion that z can be socially preferred to x can, of course, be directly derived using utilitarianism, maximin, leximin, or some utility-based equity criterion. However, we might wonder whether those approaches should be used in the case of judging torture. But the decision to rank y over x by any of these criteria in a choice involving no judgement of torture, will readily translate into a preference for torture-inclusive z over x, due to welfarism”. (reproduced from Sen).³¹

Ng has commented on this analysis.³² He states that equity (and therefore other “desirable” characteristics) can be incorporated into the welfarist view. In this view, the goods and services described by Culyer,²¹ can be extended to philosophical concepts such as equity, fairness and morality. Using the torture analogy, Ng asks why we think of torture as wrong. He argues that ultimately in answering this question, we will give a welfarist argument. Ng uses the example of someone who enjoys walking, but is unconcerned for the welfare of stones (and their right not to be stood on). We don’t care about stones’ rights because they have no feelings and thus suffer no disutility from being stood on. If stones did have feelings and therefore experienced disutility, this should be incorporated into the welfare of society. Clearly, this is a welfarist argument, though getting all the relevant utility data may be impossible.

Thus, one reason for considering the rejection of welfarism in economic evaluation is because utility information is incomplete (or individuals are unlikely to, or cannot, reveal true values) and the opportunity cost of getting this data for each decision is impossibly large. This is a practical issue rather than an objection to the theoretical principles of welfarism, but is a key concern for ensuring that research itself is cost-effective.

However, there are also theoretical problems associated with the use of welfare economics, such as the Scitovsky paradox,³³ whereby an allocation of goods A is superior to another allocation B using the Kaldor Hicks compensation test, we can also show that B is superior to A using the same test. This is due to the relative

values of the two goods changing under the new distribution.³⁴ So, for example, imagine a two person (A and B) economy with each person having one apple and one pear, both of which have the value 10 utils. The economy therefore has a total utils of 40, split equally between A and B. Now if another apple is introduced and given to A, he has 30 utils and can compensate B with up to 10 utils and they would both be better off. Clearly this would pass the Kaldor Hicks test. However, if the introduction of the apple caused the relative value of apples and pears to change such that the apples were now only worth four utils and pears worth 12, then clearly A would have 20 utils (two apples and one pear), while B would only have 16 (one apple and one pear). A movement back to the original state of the economy would therefore be better for B and he could compensate A with up to four utils and still be better off. For a more detailed exposition see Varian (1992).³⁵

2.3.1.2 Valuation of utilities in the objective function

Even in instances where utility information approaches completeness there is the question of how utilities should be valued. One way utilities can be valued is via willingness-to-pay (WTP) which is derived directly from welfare economic theory,³⁶ and in practice is the most common method of valuation of benefits in cost benefit analysis (CBA). However, in practice, WTP figures are likely to be closely related to ability to pay and that using this method to measure outcomes could result in a distribution of benefits that favour the wealthy. The aggregation of utilities expressed in monetary terms is clearly problematic when the income/wealth of individuals (and therefore the marginal utility of money) is very different (see 2.3.2.4 below). Other problems with the use of WTP are that individuals may not know their WTP or that they may not state the value that they would actually be prepared to pay. However, this problem is not exclusive to welfare economics or welfarism. WTP can (and indeed is) used in the extra-welfarist tradition (see section 2.3.2 below) as a means of valuing health.

More specific to the welfarist tradition is the issue of whose utility values should be used. As stated above, in this paradigm, individuals are the best source of their own utility information (consumer sovereignty). Thus in theory, utilities should not

be estimated by anyone other than the recipient and social decision making based on evaluations carried out in this sphere would be a huge undertaking. Indeed, it has been asserted that in the case of health, consumer sovereignty does not hold and that therefore techniques such as WTP that are reliant on this assumption are irrelevant in the case of health care decision making.¹⁹

Nevertheless, using these principles, and thereby generating a measure of societal welfare that is based on the summation of individual utilities (themselves based on individual valuations of the consequences on consuming goods and services), the welfarist approach not only allows a comparison of how resources should be allocated within the health care sector, but also allows a comparison of how resources should be allocated between health care and other activities (or sectors of the economy). This cannot be accomplished in the extra-welfarist approach described in section 2.3.2, where other characteristics such as health are seen as suitable maximands. These sector specific maximands have no substantive meaning outside the relevant sector, though trade-offs between sector specific maximands could, in theory, be incorporated into an assessment.³⁷

2.3.1.3 The Initial distribution

It is worthwhile noting that monetary valuation is not the only means of assessing utilities in welfarism. However, whichever metric is chosen, the potential Pareto improvement would require that the gainers from a change could in theory compensate the losers in the appropriate units. Those who have a greater endowment at the outset may therefore be favoured by welfarism. It is implicit within this paradigm that the initial distribution (of money, wealth, happiness etc) is at least acceptable, if not desirable. This is a value judgment that many would regard as incongruous with a “fair” system.

2.3.1.4 Other issues with welfarism

Sculpher et al³⁸ and Claxton et al³⁷ highlight the implausibility of the assumption that we live in a first best world where conditions exist for free markets to operate efficiently. Clearly, this is not the case. The economy is characterized by many imperfections, such as asymmetric information between producers and

consumers, few producers and the existence of public goods and externalities. As we live in a second best world (where free markets do *not* operate efficiently to maximize social welfare), choosing between interventions based on prices that are assumed to be derived in a first best world may actually reduce social welfare.³⁹ Other authors have highlighted the imperfections in the health system, and thus the concern with reducing social welfare by assuming a first best world may be particularly relevant for the health sector.

In practice, potential Pareto improvements are not likely to be recompensed (i.e. that winners in reality have no requirement to compensate losers). Thus, not only is utility information incomplete and difficult to value, but also transfers to recompense the losers are unlikely to occur.

The difficulties arising from using welfare economics in economic evaluation in health care has led to the development of the “extra-welfarist” approach (also known as “non-welfarist”). This approach forms the basis of most cost-effectiveness analyses, in which health (typically) rather than welfare is considered the objective to be maximised.

2.3.2 Extra-welfarism and social decision making

In the extra-welfarist paradigm, the orthodox assumptions about social welfare described above are exchanged for another set of objectives. Richardson states that most would consider the main objective of health care systems to be the maximisation of health rather than utility.⁴⁰ In Sen’s capabilities approach,²⁷ health is considered to be one of the “most important conditions of human life and a critically significant constituent of human capabilities”.⁴¹

In the social decision making approach, the problem of how best to allocate scarce resources becomes a constrained optimisation problem. There is a given social objective (for example, the maximisation of the health of society) which is exogenously defined, and a given budget constraint which is again exogenously defined. This approach does not allow assessment of how much should be spent

on health compared with other sectors, merely given this exogenously defined budget, how best should it be spent.

2.3.2.1 What is in the decision makers' objective function under extra-welfarism, and how is it valued?

In the extra-welfarist approach the contents of the objective function tend to be narrower than under welfarism, though there is no requirement that this is so. Indeed, the extra-welfarist position, largely initiated by Sen in his capabilities approach,^{20, 27} stresses the move away from the narrow concept of measuring individual utilities. Nevertheless, the analytic technique most commonly used in the extra-welfarist economic evaluation is cost-effectiveness analysis (CEA), where there is a single unidimensional outcome measure, such as the QALY, to be maximised. Most analyses conducted in the extra-welfarist paradigm (see Chapter 3 and Richardson et al 2005⁷) use an outcome measure (such as the QALY or more narrowly a change in blood pressure or Body Mass Index) where it is assumed that there is no difference between individuals in the level of utility derived from a unit gain in outcome. That is, the value of a QALY is independent of the marginal utility of money, or that any differences between individuals are not relevant to the decision. In this “unadjusted” state, all QALYs (or change in blood pressure or whatever outcome is being assessed) are valued equally and therefore differential ability to value improvements is avoided.

The use of a composite measure of the quality and quantity of life has become more common in the recent literature, in particular with the use of the QALY.¹⁹ At this point it is worth commenting on different “types” of QALYs. Although QALYs are always the product of the time spent in a specific health state and the “quality” of that health state, the derivation of this quality weight has different potential sources. Hurley¹⁹ identifies two broad types of QALY, those where the weights are derived using psychometric testing (see for example Weinstein and Stason⁴²), and those where the weights are derived using “utility” weights. In the former case, individuals are often asked to rate states on a scale of death to perfect health (and there is therefore no allowance for uncertainty), while in the latter individuals are asked to trade-off hypothetical states (where uncertainty is incorporated into the decision). Most QALYs used in the literature are “utility”

based QALYs,¹⁹ though it does not follow that there is consensus as to the meaning of a QALY. To add to the confusion, Richardson and Manca⁴³ found that QALYs were also calculated in different ways with a variety of assumptions, though later work suggests the “correct” methodology for QALY calculation.⁴⁴

Some authors argue²¹ that the QALY represents a subjective assessment of health *not* utility, even though utility theory is used in the construction of this QALY, while others contend that the QALY is a measure of individual utility.⁴⁵ Hurley¹⁹ points out that therefore the use of a utility based QALY can be used in either a welfarist or extra-welfarist analysis, and that the use of this instrument does not identify the paradigm chosen; the same author also states that the use of the utility based QALY is increasingly popular in economic evaluation.

Other authors^{46, 47} have argued for “equity weighted” QALYs, where certain beneficiaries of QALYs would be valued higher than others (for example, the young may be favoured above the old). Other authors have argued against “consequentialism” in economic evaluation and that the inclusion of procedural preferences is important.^{48, 49}

While the contention of the above authors is that QALYs should be measured but enhanced in some manner, other authors have cautioned against the use of QALYs at present due to the strong assumptions required in their generation,⁵⁰ while McGregor (2003) considers that shortcomings in the methodologies used to generate QALYs warrants great care in their use.⁵¹

Others authors dispute the usefulness of QALYs at all,⁵² though the author appears primarily concerned with their use when there is no alternative treatment available for that patient group. The implications of this approach would be to have a higher cost-effectiveness threshold for conditions where there is no existing treatment, and presumably the denial of treatment (given a fixed budget) to some patients with complaints that are currently treatable, even though these patients would benefit more from treatment. This is an individual value judgment but even if the author’s argument were accepted, there is no alternative mechanism to assess by how much the cost-effectiveness threshold should be raised. In principle a higher weight could be attached to the gains from treating patients

where there is no existing treatment (the argument takes a similar path to that of the “rule of rescue”, where identifiable individuals facing death are deemed a treatment imperative).

Advocates of the “QALY” approach argue that the limitations of the QALY are well known, and that the explicit assumptions prohibit distortion of results from those with a vested interest.⁵³

2.3.2.2 Expanding the objective function of the decision maker

In the extra-welfarist approach, if health is the only consideration in the objective function, the only means of increasing social welfare is through increasing the health of society. The existence of such an objective function necessitates an entity with a legitimate claim to make such decisions³⁷ based on this function. In addition, the methodology used in analyses should be explicit about the social and scientific values that are made; thus the *legitimate* decision maker can be seen to be making *legitimate* decisions. The form of this entity is discussed in section 2.3.2.3, where democratically elected governments and their appointed agents, are considered potentially appropriate decision making entities.

The objective function of the decision making entity, and therefore the extra-welfarist approach, does not preclude a broader perspective than simply health. However, if health is the starting point, then to include other items in the objective function requires knowledge of the rate of substitution between health and these other items.³⁸ As a crude example, if the decision maker has individuals’ autonomy as a desirable outcome *per se*, he/she would need to have some valuation of autonomy in terms of a QALY (e.g. a point increase in autonomy equals 0.1 QALY), where QALYs are the measure of health. An example of the potential expansion of the QALY as an outcome measure is considered in Chapter 7.

However, if characteristics other than health are in the objective function and are included as “benefits” or “costs” in the CEA, it is unclear how these should be viewed. For example, consider two interventions A and B with identical costs.

Intervention A generates 3 QALYs of health benefits and 2 QALYs (equivalents) of non-health benefits (assuming we know the rates of substitution and can therefore calculate QALY equivalents). Intervention B generates 4 QALYs of health benefits but zero non-health benefits. Clearly, intervention A is “better” in that it produces more QALY equivalents for a given sum, but if health is the maximand, B would be chosen. This topic has, as yet, received little attention in the literature, but is likely to generate more interest as the emphasis on broader outcome measures and cross-sectoral effects increases. However, the least that is required is that the perspective of the study, the exact nature of “relevant” evidence, and the constituents of the objective function must be stated *a priori*.

2.3.2.3 Inclusion of other “socially desirable” characteristics in the extra-welfarist paradigm

The section above describes how other outcome measures can be incorporated into the extra-welfarist paradigm. While it may be feasible to obtain the MRS between health related outcomes and, for example, equity, this is a time consuming and conceptually difficult exercise.

Others have argued that we use, as a simplifying heuristic, a series of rights/rules,³² that are perceived to improve social welfare (though it should be noted that Ng used this in the welfarist paradigm). Nozick⁵⁴ (1974) has argued that these rules “do not determine social ordering but instead set the constraints within which a social choice is to be made”. Government could be seen as the overseer of these rights and also an appropriate decision making body. The two roles of government are likely to be related; in, for example, providing a NHS, the government gives individuals in that society access to a minimum level of care, in providing schools, government gives individuals (the opportunity of) a minimum level of education and so on. This provision of goods/services may be seen as rights, and the level of these rights may in part determine the level of funding for each service. The legitimacy of government to act in this manner is gained, as argued by Sugden and Williams,²² through the political process.

Sculpher et al³⁸ (2005) suggest that there are institutions in the UK that have been given authority to make decisions (about health care) on behalf of society as a whole. These authors go on to suggest that institutions such as the National Institute for Health and Clinical Excellence (NICE) provide prescriptive recommendations based on explicit or transparent assumptions, and that this is preferable to the welfarist approach where many of the assumptions remain implicit.

Thus in the extra-welfarist world, a legitimate external body is seen as the overseer of individuals' rights and freedoms. Subsequently, this external body delegate responsibility to a (health care) decision maker who has a remit to produce the maximum amount of health for a given cost. It is argued that the political process delivers governments that act in a manner consistent with this external body and that this process gives the government a legitimacy to behave in this manner. Critics of the extra-welfarist approach³⁶ contend that this approach simply substitutes the preferences of the decision maker or the community for those of individuals. While this is a valid criticism, the authors do not present an alternative that adequately deals with the inequity of initial distributions, as discussed in section 2.3.2.4 below.

2.3.2.4 The Initial distribution in extra-welfarism

As described above, the welfarist paradigm is not value free when considering distributional issues. In particular there is the controversial assumption that the initial distribution (of utility, however defined) is acceptable.

In the extra-welfarist paradigm, there is no implicit assumption that the initial distribution is acceptable. One of the founding principles of the UK NHS was that access to health care should be based on capacity to benefit rather than the ability to pay and that there should not be differential treatment because of income.⁵⁵ Thus it would seem reasonable to use a methodology (such as CEA) which makes no *a priori* assumptions regarding the acceptability of the initial distribution of goods/services to maximise a social objective, rather than a methodology (such as CBA) where there is an implicit assumption that the initial distribution is acceptable

thereby generating results that are likely to favour those with a large initial endowment.

2.3.3 Concerns with the extra-welfarist approach.

Sheldon⁵⁶ has expressed concern that economic evaluation places too much emphasis on health rather than on welfare. Factors other than health may be of value to society and may be appropriately considered in economic evaluation. However, as argued above, the use of CEA with the QALY as the measure of benefit does not preclude the inclusion of other characteristics which may be important to society. Attempts have been made in the discrete choice experiment literature to value these other important characteristics, though none have, as yet, transferred these other characteristics into the QALY framework. In Chapter 7 the identification of these characteristics and how they might be incorporated into cost-effectiveness analysis is discussed more fully.

There is also the question of what constitutes “health”, assuming that health is a legitimate objective function for health care decision making. Some would argue that psychological outcomes that are not picked up elsewhere in HRQoL, should be included in health outcomes. Even if these outcomes are not considered to be included as “health”, they may still be an appropriate part of a decision makers objective function.⁵⁷ A further problem may be the lack of meaning of the health maximand outside the field of health (that is as a general maximand).¹⁷

Nevertheless, this does not preclude extra-welfarist analyses in other spheres,¹⁷ and potentially the synthesis of costs and effects across different sectors.³⁷

However, the desire to broaden the perspective from simply health requires that the Marginal Rates of Substitution (MRS) between health and the other characteristics to be included in the objective function are known. The discrete choice experiment described in Chapter 7 elicits these values for patients with chronic conditions enabling the relative values of health and other “important” outcomes to be ascertained. In principle, this methodology could be used to include other “socially desirable” characteristics, such as equity.

Birch and Donaldson³⁶ are critical of the move away from welfare economics in the field of economic evaluation. They consider utility of the individual to be of paramount importance, rather than “health” and argue that “utility” can be broadened to encompass concepts such as utility derived from processes in the welfarist paradigm. The authors neatly encapsulate the difference between the paradigms using the example of smoking. In the extra-welfarist paradigm where health has been designated as the only outcome we are interested in, a programme that cuts smoking levels would be judged as “beneficial”. However, in the welfarist world, individuals’ utility may be lower through stopping smoking, and it is less clear whether the programme would be beneficial or not.

The same authors are also critical of the potential paternalism within extra-welfarism, where the individuals’ assessment of his/her own welfare (individual sovereignty) is replaced by that of a decision maker of the general population. The objective of individual sovereignty has itself been questioned. For example, Sen⁵⁸ comments that individuals with disabilities can adjust their expectations down as they adapt to the disability. Hence, as utility is in part determined by the individual’s assessment of the situation relative to their expectations, they may have high levels of utility.

2.3.4 Welfarism, extra-welfarism and the economic evaluation of interventions to support self care

Does the choice of philosophical paradigm impact on the conduct, analysis, results, conclusions and recommendations of an economic evaluation? Without conducting analyses in both paradigms, it is not possible to say whether results and conclusions and therefore recommendations would be affected. The conduct of the economic evaluation would need to be very different. Consider the outcome measure in the study, the extra-welfarist perspective allows the measurement of a variety of outcome measures, some more useful than others, as a measure of health. These measures are often biological measures or patient responses to a question (or series of questions) at various time points. This measure can then be expanded if required. The welfarist perspective however, requires individual utilities to be measured. This would require the impact of interventions on health and all other factors (and interactions) that may impact on individuals’ utility. This

is a more complex task that would impact on the cost and duration of economic evaluations.

2.3.5 Discussion of normative foundations of economic evaluation

So, what are the differences between the two approaches? Brouwer et al¹⁷ identify four principal differences. These are:

- i) the outcomes considered. Welfarism uses the utility of individuals, extra-welfarism *can* include other outcomes, though in practise, health is commonly the maximand.
- ii) the sources of the valuation of outcomes. Welfarism uses the individual, extra-welfarism can be the individual but can also be a decision maker or the general public.
- iii) how the outcomes are weighted. Weighting is not allowed under strict welfare economics criteria (as it is not certain that weightings are actually utility information) but are an integral part of extra-welfarism with the use of equity weighted QALYs for example.
- iv) interpersonal comparison of outcomes. Again this is not allowed in the strictest welfare economics interpretation, but is incorporated in extra-welfarism as *ceteris paribus*, an intervention generating 2 QALYs for individuals A and B is considered superior to an intervention generating one QALY for individual C.

The similarities between the practical implementation of both approaches (which tend to be cost-effectiveness analysis in extra-welfarism and cost-benefit analysis in welfarism) has also been the subject of debate with some authors stressing the similarities,⁵⁹ while others dispute there is any link at all.⁶⁰

There are drawbacks to both the extra-welfarist and welfarist paradigms. In the extra-welfarist approach, the absence of a legitimate decision maker could be considered problematic. However, the politically elected government can be considered to have been given a mandate to set budgets for sectors,⁶¹ as well as maintaining the rights and freedom of individuals.⁵⁴

There are also significant problems with the use of welfarist economics in economic evaluation. The focus on individual utility and the incompleteness of this

utility data, as well problems valuing these data, are practical issues in the use of welfarism in economic evaluation.

Choices need to be made between paradigms. The pragmatic advantages of the extra-welfarist approach, using a commonly used, understood and simple outcome such as the QALY are judged to outweigh the disadvantages. For the purposes of the following analyses, the constrained optimisation approach of extra-welfarism is preferable, so long as the perspective is wide enough. The implication for economic evaluation of interventions to support self care is that, though there are limitations to the use of the extra-welfarist approach in assessing the cost-effectiveness of interventions to support self care, for practical purposes this is the appropriate paradigm.

2.4 Single trial based economic evaluation and decision making

The choice between the extra-welfarist and welfarist approaches is not the only ongoing controversy in the field of economic evaluation. Economic evaluation within randomised controlled trials has been seen as the primary source of evidence of cost-effectiveness for decision making. However, just as the clinical evaluation field has moved towards the incorporation of other relevant information, using techniques such as meta-analysis, so the field of economic evaluation has shifted towards the synthesis of different forms of relevant evidence from sources outside the RCT.

Most economic evaluations of interventions to support patients' self care have been conducted in the extra-welfarist tradition described above.⁷ In addition, the majority of these economic evaluations were all based on a single trial. Where trial data were available, there was no attempt to use data from other sources. This reflects the historical use of economic evaluation alongside clinical trials as the primary means of assessing the cost-effectiveness of many interventions⁶² including those in the self care sphere⁷.

Recently it has been argued⁶² that, in most cases, the use of "a single RCT as a vehicle for economic analysis will be an inadequate and partial basis for decision

making”, though it is acknowledged that RCTs are useful in providing an unbiased estimate of the relative treatment effects of comparators and ensuring that differences are due to the intervention. The authors consider that these single trial based economic evaluations have four important drawbacks that limit their usefulness to decision makers. Firstly, they fail to consider all the relevant alternatives. Secondly the time horizon of the study rarely captures all the outcomes of interest. Thirdly, trial based economic analyses are not always relevant to the decision context and finally, these analyses do not consider all the relevant evidence.

While these are persuasive arguments, in certain cases there remains a case for presenting a single trial based analysis; it is asserted that in the evaluation of the EPP, presentation of a single trial based analysis is one appropriate method. Each of the criticisms of trial based economic evaluation is addressed below.

2.4.1 Failure to consider all relevant alternatives

The Expert Patients Programme (EPP) based on the Chronic Disease Self Management Programme (CDSMP) is probably the only group therapy aimed at improving self-efficacy and ultimately health status, delivered to patients with a range of chronic conditions in the UK. Other self care support interventions have been disease specific (see the literature review in Chapter 3) or do not include self-efficacy as the mode of improvement. Therefore, it can be argued that in this instance where there are no obvious alternative treatments or management options for this group of patients, evaluations that consider CDSMP should be considered.

In addition, the inclusion of all relevant alternatives necessitates the inclusion of indirect comparisons. In principle, treatment as usual could be considered a common comparator and links made via this treatment option. However, in this instance, treatment as usual may not be a constant treatment. There are differences in what constitutes treatment as usual with different clinical practices in different countries with different patient populations and a variety of conditions. It

is therefore a moot point as to whether treatment as usual can be used as a common link.

2.4.2. Truncated time horizon

Economic evaluations conducted within or alongside clinical trials are often limited to explore the costs and effects within the time frame of the clinical follow-up period. This may not be long enough to capture all the relevant costs and benefits of interventions and the use of such a truncated time horizon can lead to misleading conclusions.^{62, 63} In the example of the evaluation of the EPP presented in Chapter 4, patients were followed up six months after their randomisation to the intervention or the waiting list control. This may not capture all the effects of the intervention. The standard “within trial” estimate of cost-effectiveness would only consider costs and effects occurring within this period. This can lead to erroneous conclusions. However, simple modelling allows the extrapolation of effects of the intervention to periods of greater than one year. The effect of treatment on costs and effects after one year is unknown. Where an intervention provides, for example, additional survival at one year, it may be reasonable to assume that the intervention has no further impact and that the additional survival at one year generates additional QALY gains over the next 15 years.⁶³ However, this assumption (effectively that survival curves are parallel after one year), may not be reasonable for a behavioural intervention where the effects may dissipate over time,⁶⁴ or where the durability of the effect is unknown.⁶⁵ Indeed, the effect on resource use after six months is also unknown and could either make the intervention more or less cost-effective. Thus, a time horizon six months to one year is deemed appropriate, though it is acknowledged that any incremental costs and/or benefits of the intervention after this time period will not be included.

2.4.3. Lack of relevance to decision context

Sculpher et al comment that the patient population considered in RCTs may be of limited comparability with the patient population being considered by the decision maker.⁶² This argument can be taken to either extreme. It could be argued that to make the evidence most relevant to a decision context, a RCT in the same

population in every respect, with the same condition, using the same intervention is the only evidence relevant to that decision. In the case of the EPP evaluation, the study was driven by the decision context. The EPP evaluation was based on patients in the UK only, and as the jurisdiction making the decision is also UK based, the decision context is appropriate. The trial was multi-centred and designed to maximise external validity with patients from a wide variety of Strategic Health Authorities, so that the evaluation should be generalisable to the patients eligible to receive the EPP.⁶⁶

2.4.4. Failure to incorporate all relevant evidence

The most fundamental criticism of economic evaluation conducted alongside a single trial is that not all the relevant evidence is included. Other trial evidence may be available estimating the relative treatment effects, but importantly, non trial based evidence may inform non treatment effect parameters. For example, in evaluating drugs for heart disease, a single trial may have a limited amount of evidence of quality of life after myocardial infarction, but there may be rich sources of data available elsewhere. Ignoring these data, it is argued, leads to “a partial or potentially misleading” analysis.⁶²

The identification of relevant evidence and the criteria for determining relevance is a constant theme in this thesis. It is argued in a later chapter that there is no simple dichotomy between relevance and irrelevance. The concepts lie on a continuum, with some evidence/data/studies more relevant than others. The incorporation of additional evidence that is deemed relevant requires additional assumptions about the relationship between the treatment effect and costs and QALYs, and the mechanism by which costs and QALYs are affected. In essence, this is a debate about the structure of the model. Each representation of the cost-effectiveness of the CDSMP, from the single trial based analysis described in Chapter 4, to the evidence synthesis presented in Chapter 6, is a model. These models use different data and require different assumptions, but *a priori* we do not know which model is “correct”. We therefore have a problem with “structural uncertainty”. We do not know all the mechanisms by which the treatment impacts on outcomes nor whether these would be reproducible in future experiments. For

example, there may be a large number of intermediate or surrogate outcomes through which the final outcome is mediated. It may not be possible to model all these mechanisms (or indeed we may not be *aware* of all these mechanisms), or it may be that differences are due to truly random variation. There is a small but developing literature that considers the subject of the “structural uncertainty” in the development of economic models.⁶⁷ Structural uncertainty has been largely ignored as methodologists have concentrated on addressing parameter uncertainty⁶⁸, which often implicitly assumes that the model structure is known. Few papers discuss the issue, and fewer formally address the issue in their analysis.⁶⁷ Techniques such as scenario analysis have been used on occasion in the health economics literature.⁶⁹ Scenario analysis examines the impact on costs and outcomes of different scenarios (for example considering a male vs female population or a younger vs older population). Other techniques such as model averaging have yet to be used in the health economics field.⁷⁰ However, even where structural uncertainty is acknowledged, and “captured” within the model, this can only be conducted when we are aware that we are uncertain. It is an empirical question, which is addressed partly in Chapter 6, whether assuming that model structures that we are not aware of simply reflect “noise” influences the results of models.

In the example used in Chapter 4, the single trial based evaluation of the EPP is essentially a scenario analysis where all other data are deemed irrelevant. This model is extended in later chapters to incorporate additional evidence with different degrees of relevance. Is it reasonable to treat one single RCT as providing the only evidence? In this instance, there are reasons for suggesting that this *maybe* appropriate. Many other trials of CDSMP are based outside the UK,⁷¹⁻⁸² and/or use a limited study design.^{71, 75}

It is reasonable to argue that only UK studies that establish the effectiveness of CDSMP using an RCT design should be included. However, there is more than one trial based analysis that meets this criterion. One other trial considered the use of the CDSMP in the UK, but this study was based in a particular ethnic patient group, rather than the general population.⁸³ Though this study was not in a nationally representative population it is acknowledged that this study might be

considered relevant, and for this reason synthesis of this trial with the EPP evaluation is conducted in Chapter 5.

More data from different locations in different settings with heterogeneous patient groups is added in Chapter 6. While this undoubtedly includes more evidence, the relevance to the UK decision context is lower than that of the single trial evidence. The implications of synthesising these data are discussed in terms of the impact on the decision and the characterisation of uncertainty.

2.4.5 Discussion of single trial based evaluations

The concerns with the use of single trial based economic evaluations are well founded. Indeed, the implications of the above critique is that, in most cases, single trial based economic evaluation is an incomplete analysis and should be supplemented with all other relevant data. The impact that expanding the evidence base could have on the decision problem should be clear. The results of one trial may well present a positive assessment of the cost-effectiveness of an intervention. However, the inclusion of other evidence that does not support the cost-effectiveness of the intervention could impact on the results and recommendations. Nevertheless, in certain circumstances, single trial based analysis remains *an* appropriate methodology (though not *the only* appropriate methodology). In the case of the single trial based economic evaluation presented in Chapter 4, there are no obvious alternatives omitted, the relevant time horizon for such studies is ambiguous, and the decision context was the rationale for the trial. The inclusion of all relevant evidence is more contentious. Relevance is a subjective concept. The question then becomes “who defines relevance?” Pragmatically, relevance could be used to justify exclusion of any study not carried out in the jurisdiction of interest, or examining the population of interest, or does not use a design that is deemed appropriate. For the purposes of the evaluation of the EPP a *starting point* for the analysis is that only UK based trials in the general population are relevant to the decision problem stated in section 1.2. Later chapters expand the evidence base to include additional UK and non-UK evidence and examine the impact of their inclusion on results and conclusions.

Therefore, in the case of the evaluation of the EPP, the presentation of results based on the single trial is appropriate, while the possibility of synthesising these data with other available data, and the impact this would have on results and decisions, is considered in later chapters.

2.5 Statistical analysis and interpretation in economic evaluation

2.5.1 Bayesians and Frequentists

Statistical techniques used in the analysis of economic evaluations are another topic on which analysts are divided. Bayesians hold that probabilities represent the degree of belief in an event occurring. This probability can be updated as and when further information becomes available. This contrasts with frequentists who only consider probabilities as the relative frequency of events (from the sample space of all events) occurring in a set of random experiments. Where an experiment cannot be repeated numerous times, a frequentist could not assign a “probability” of an event occurring as the long run frequency of an event is seen as its probability (hence the term “frequentist”)

For CEA, the major feature of Bayesian thinking is that parameters are considered to have a distribution, rather than be considered as a single “true” figure as in the frequentist paradigm.⁸⁴ Thus for Bayesians, some values of parameters are more credible than others based on both the “data *and* prior beliefs”.⁸⁵ Hence, in Bayesian CEA, we estimate a posterior distribution (for example of the population mean cost) using the available data but adjusted by our “prior” beliefs, which may in fact be other data from other trials, observational evidence, or other form of evidence. These “priors” are intended to express the beliefs we hold before we have had chance to observe any data. Often these “prior” beliefs are given very little weight so that the accumulation of a large amount of data renders them extraneous.

Bayesians can talk about the “probability” of an intervention being cost-effective at some cut-off point given the data. Frequentists cannot make such statements. It has been argued that the Bayesian approach when used in economic evaluation

gives a natural framework for the assessment of uncertainty in decision making and that results are more useful to decision makers.^{86, 87} When the threshold value of the outcome of choice (for example the QALY) is known, Bayesian analysis can be used to estimate the probability that the intervention is cost-effective, that is to estimate the probability that the cost per QALY is below a given threshold. The ICER can still be presented as a measure of the relative efficiency of an intervention, while the probability that an intervention is cost-effective, at various threshold values of a QALY, can be used as a measure of uncertainty around the decision and the potential benefit from conducting further research.^{88, 89} The uncertainty around the decision is often portrayed graphically using the cost-effectiveness acceptability curve (CEAC),^{88, 90} while value of information techniques are often employed to value future research.^{91, 92} As the Bayesian approach can incorporate new evidence as it becomes available and allows presentation of results in a format that is more useful for decision makers, it is considered the appropriate paradigm to conduct economic evaluation.

2.5.2 The role of inference

Statistical inference can be used in both the Bayesian and frequentist paradigms. However, it is in the frequentist paradigm that inference and the use of statistical significance has been most prevalent. This involves setting up a hypothesis and then collecting data, (for example, from an experiment) to inform this question and subsequently either accepting or rejecting the hypothesis based on a previously stated “level of statistical significance”. Statistical significance levels of 5% are commonly employed in the literature, though they are commonly misunderstood or misinterpreted (see for example Altman and Bland).⁹³ Other authors have pointed out the problems associated with significance testing and inference (either frequentist or Bayesian) and decision making.⁹⁴ Claxton points out that if we accept a null hypothesis (of no difference) when one treatment has a higher (but not significantly higher) expected net benefit, there is a cost implication of not implementing the intervention that can be measured in terms of benefits forgone by those who are not treated.

As an example, consider an economic evaluation of two interventions for the treatment of a chronic illness. The economic analysis may not be able to demonstrate a “statistically significant” improvement in patient outcomes, nor a “statistically significant” reduction in costs (possibly due to the study not having sufficient statistical power). In this instance, classical statistical inference fails to reject the null hypotheses that there is no difference in effectiveness or costs between two groups, and makes few recommendations for decision makers.

In contrast comparing the incremental costs and effectiveness of an intervention, regardless of whether these differences are statistically significant or not, generates an estimated cost per QALY. Assuming that we have a decision maker with legitimacy (see section 2.3 above), and an exogenously defined budget constraint (which may come from this decision maker or another with legitimacy to allocate budgets between sectors), then the decision maker can compare the cost-effectiveness of the intervention with the threshold cost per QALY. This approach does not necessitate that these differences in cost or effect are “statistically significant”, though it is common practice for analysts to present confidence intervals around the difference in means, from which statistical significance (or lack thereof) can be ascertained.

Clearly there are implications of accepting one paradigm over the other and the different statistical techniques can yield different results.⁹⁵ If classical statistical inference is the paradigm of choice, then fewer analyses will be informative for decision making purposes as the requirement that differences are “statistically significant” is more restrictive than demonstrating the probability of an intervention being cost-effective. To be informative for decision making, it should not be a requirement that incremental costs or effects (or both) are statistically significant at some arbitrarily defined threshold. Incremental costs and effects can be provided together with estimates of uncertainty and estimates of the probability that interventions are cost-effective given a threshold value for the outcome measure. The fact that these estimates can be provided over a range of threshold values means that the analysis is not confined to pre-defined and arbitrary levels of statistical significance. For these reasons, the use of the either classical

statistical inference or its Bayesian equivalent are rejected in the interpretation of results of economic evaluations.

2.5.3 Simulation in Bayesian cost-effectiveness analysis

Bayesian methods are becoming increasingly used in CEA.⁸⁴ The posterior distribution, where both the data and other external evidence (or beliefs) are incorporated should be used to inform decision making. In some simple cases, this posterior distribution can be estimated using standard notions of probability. For example, the probability of tossing eight heads from ten tosses of a fair coin given that the first two tosses were heads, can be relatively simply annotated and calculated. Alternatively, we could write a simple programme to toss a coin ten times with the first two as heads. In this instance, our prior belief may be simply that the coin is fair and that there is therefore a probability of 0.5 of either a head or a tail on each toss. Replicating this procedure, say, 20,000 times would yield a similar estimate to the standard method above. While the former approach may be more appropriate when the solution is tractable (as in this example), for more complex problems without tractable “closed form” solutions, simulation is a suitable alternative. The vast majority of problems in CEA would not be tractable (or at least it may be necessary to make unrealistic assumptions concerning the distributional forms of priors so that they are “from the same family” as those of the likelihood/data) and simulation is a means of estimating the posterior distribution and thereby enabling a useful Bayesian solution.⁹⁶

Markov Chain Monte Carlo (MCMC) is one simulation method that has proved particularly useful. WinBUGS is a freely available software programme that implements MCMC using Gibbs sampling (BUGS is an acronym for Bayesian inference Using Gibbs Sampling).⁹⁷ To simulate the posterior distributions (in effect, the final estimates, and the distributions, of the parameters of interest) of random variables, we first choose an initial value for each parameter. Then, based on this initial value, the conditional distributions we have defined and our prior expectations, we draw the next value.⁹⁸ Using this value and the conditional distributions, we draw the next value and so on until (hopefully) there is convergence, where continued sampling results in a stationary distribution.

Checks for convergence and other model diagnostics are an important consideration in these analyses, and they are described in more detail in a later chapter.

2.6 Discussion

This chapter has considered the schools of thought within economic evaluation, particularly the evaluative space (whether we consider individual's utility as the only contributor to social welfare), what evidence can/should be included in an economic evaluation and the statistical paradigm that should be implemented. Choices need to be made regarding the paradigms of choice, but those conducting analyses should be aware that these choices can impact on the results and recommendations for the decision problem identified in the first chapter. This chapter has demonstrated that though there are conflicting paradigms within which economic evaluation could be based, the extra-welfarist position with health as the maximand can be justified so long as there is an agency with legitimacy to make decisions to maximise health. It is argued that the democratically elected government forms a suitable agency and that bodies such as NICE, which are empowered from the democratically elected government, have a mandate to make decisions to maximise society's health. In this extra-welfarist perspective, societal health rather than (the aggregation of) individual utility is considered the maximand. However, the original motivation for extra-welfarism was to expand the outcome measure of interest. The use of health as a maximand could be considered a narrower maximand than utility, therefore expanding the outcome measure beyond that of simply health is a justifiable objective. This is considered in Chapter 7 where the relative value of health to other important factors is assessed and included in the decision problem.

The issues around the usefulness of trial based evaluation have also been discussed. It is acknowledged that a single trial based analysis is only appropriate in certain circumstances. However, it is plausible that these circumstances are applicable in this instance and that a single trial based analysis may be the appropriate model. Therefore, a single trial based economic evaluation of the Expert Patients Programme is presented in Chapter 4. Extensions and variations

on this single trial based model are presented in Chapters 5 and 6. These chapters introduce additional information, firstly from another UK based evaluation (in Chapter 5), and then through the incorporation of non-UK based data (in Chapter 6).

It has also been argued that the Bayesian statistical paradigm is appropriate for the analysis of economic evaluations and that the role of inference (either Bayesian or classical) is inappropriate for economic analysis.

The choice of paradigms for the analysis and interpretation of the cost-effectiveness of one particular intervention to support self care has been established. However, there are many interventions designed to support self care, some of which have been subjected to economic evaluation. Do the existing economic evaluations of self care support conform to the above arguments? Do economic evaluations of self care support interventions use methods of evidence synthesis to incorporate “relevant evidence”? Do they use an outcome measure that is wider than simply health (or a subsection thereof)? Do they assume a Bayesian or frequentist approach and is decision making based on concepts of statistical inference?

These issues are often implicitly included in checklists used to assess the quality of existing evaluations. For example, the use of relevant evidence and the use of a narrow outcome measure are included in the Drummond checklist,⁹⁹ via the questions “were all important and relevant costs and consequences identified?”

So, the paradigms chosen are one aspect of quality assessment. Clearly, there are other considerations, such as the use of an appropriate perspective, the use of discounting and the discussion of generalisability. How all of these issues are dealt with in the existing literature of economic evaluations of self care support interventions is discussed in the systematic review presented in the next chapter.

2.7 Conclusion

It has been argued in this chapter that there are several contentious issues in the design, conduct and analysis of economic evaluation. For the purposes of this

thesis, analyses use a Bayesian approach to statistical analysis, and employ an extra-welfarist perspective. The issue of single trial based evaluations compared with evaluations including more evidence is examined in detail in later chapters.

Chapter 3. Systematic review of economic evaluations of self care support

3.1 Introduction

Before considering new research, it is necessary to consider work that has already been published. Existing literature may address the decision problem, but may also inform the design, conduct and analysis of future research. This chapter examines the existing literature around interventions to support self care and assesses whether this literature can be used to inform the design conduct and analysis of the cost-effectiveness study presented in the following chapter. An earlier version of the review of published economic evaluations contained in this chapter has been published in the International Journal for Technology Assessment in Health Care.⁷ This chapter presents an updated and extended version of that publication.

3.2 Evaluating self care support

There are well established methods for the conduct of economic evaluations and checklists have been designed to assess the quality of these interventions.⁹⁹ It has been argued that there are additional problems in conducting economic evaluations of self care support interventions.^{7, 100} These additional problems are described below, and inform the checklist developed in section 3.3.4.

3.2.1 Comparator intervention

Economic evaluation of any intervention requires a comparator.⁹⁹ Unfortunately, the comparator is often less well defined in self care support interventions than in other interventions where the comparator is a placebo or another intervention. Historically, the usual comparator in self care support interventions is no active intervention or usual treatment.¹⁰¹ The results section below indicates that this is true also of economic evaluations in this sphere. This makes it more difficult to compare interventions against each other and hence to choose the best option. Comparing different interventions would require indirect comparisons with a

common link. In this instance, “usual treatment” could provide a common link, but there is little to suggest that usual care is consistent across disease groups let alone geographical regions.

3.2.2 Placebo effects

Outcomes can be affected by patient expectations and beliefs, which can lead to effects where at least part of any change in outcomes is attributable to there being *any* active intervention.¹⁰² Where these effects are potentially present, it is better to compare active interventions against each other rather than against a passive usual care control group. This is particularly problematic in self care support evaluations as the comparator is usually inactive.

In addition, studies of group based self care support interventions, do not allow for the potentially beneficial effect of being a member of a group of individuals with similar conditions.

3.2.3 Control group contamination

There may be a higher risk of control group contamination.⁷ Access to the intervention is more difficult to prevent with self care support interventions. For example, with patient education materials, patients in the control group may get access to the materials if they are published or from members of the intervention group. Randomisation by centre rather than by individual can reduce the risk of contamination.

3.2.4 Range of outcomes

Self care support interventions are likely to have a wider range of outcomes since they are often intended both to improve patient health and to empower patients by giving them greater control of health affecting decisions. Hence evaluations tend to include a wider range of outcome measures. For example, the evaluation of the Expert Patients Programme described in Chapter 4 includes measures of self-efficacy to manage disease, a measure of communication with health

professionals, and subjective well being in a number of domains, in addition to a battery of more conventional physical and mental health measures.¹⁰³

3.2.5 Patient costs

Evaluations ideally should adopt a societal perspective and take account of costs wherever they fall, including on patients.⁹⁹ If some costs or benefits are excluded, misleading analyses can be conducted. For example, an intervention may reduce costs slightly in the health care sector, and taking an NHS perspective would show a cost saving. In reality however, these costs may be just shifted to another provider, such as community care, so that the real cost to society of a new intervention is greater.

Patient costs, which include time off work, out-of-pocket expenses, and travel are often difficult to measure and for this reason are often omitted from analyses. Given that self care support interventions are designed to alter the way patients manage their conditions, patient costs are likely to be more important than in economic evaluations of more conventional interventions.

The inclusion/exclusion of these costs in the single trial based economic evaluation is considered in sections 4.5.4 and 4.6.7 of the next chapter.

3.2.6 Length of follow-up

There is some evidence that beneficial effects of interventions last for quite short periods.^{64, 65, 104, 105} Hence results from studies with short follow-up periods may not be reliable guides to long-term effects. Many self care support interventions involve patients with chronic conditions so that long-term consequences may be important.

There are fundamental trade-offs in choosing the length of follow-up in trials of self care support interventions. Longer follow-up may provide information on the time path of effects but it increases the risk of control group contamination. In the absence of any barriers, control group patients may employ the same self care

techniques as the intervention group, thereby diluting the treatment effect. Thus a reduction in the difference in observed outcomes between control and intervention groups over time may reflect a genuine reduction in the effect of the intervention on the intervention group or it may reflect an improvement in the condition of members of the control group who adopt the same self care practices as the intervention group.

3.2.7 Transferability of results

The transferability of results to other settings is important. Though transferability of estimates of effectiveness is an issue, in studies of conventional interventions issues of transferability are most usually raised for cost estimates. Differences in unit costs of resources between countries or over time mean that disaggregated data (volume and unit costs for different types of resource) are more likely to be useful in other settings. Differences in unit costs across settings may imply that cost minimising input mixes may differ and so a simple recalculation of costs using the original study volume data but local unit costs can be misleading.

Transferability of the treatment effect on outcomes may be more of an issue for self care support interventions than several other types of intervention. Because the interventions tend to be psychosocial in nature, cultural factors that affect patients' receptiveness to self care may influence both the no intervention baseline and the effect of an intervention. Thus it is important that the context of the intervention is clearly specified.

The transferability or exchangeability assumptions made in economic evaluations are discussed in much greater detail in Chapter 6.

3.3 Methods

3.3.1 Self care support

For the purposes of this review, self care support was defined as interventions *facilitating* patients taking decisions intended to alter the effect of their conditions

on their health, via their responses to symptoms, or monitoring their condition, or self treatment.⁷ The definition is broad and the range of interventions is wide, but the emphasis is on informing, educating or training patients to change their behaviour.

3.3.2 Inclusion criteria

Studies were included if they were considered to be full economic evaluations of self care support interventions, that is the intervention was compared with an appropriate alternative and the costs and effects of both interventions were considered.⁹⁹ Hence, study type was not restricted to randomised controlled trials.

Interventions included range from facilitated education programmes for diabetes patients to interventions supporting self care by introducing patient requested consultations rather than standard follow-up at the consultants' recommendation. Studies of interventions where the condition is managed by the patient but did not involve trying to change patient behaviour were excluded. Thus, an intervention to enhance an asthmatic patients' ability to manage their own condition by providing education and support from trained personnel would be included in the review, but an intervention comparing asthma drug A and asthma drug B would be excluded.

3.3.3 Search strategy

A systematic literature search and review of published economic evaluations of self care support interventions was carried based on the above definition. Full details are provided in Appendix A. The search strategy is identical to that previously published,⁷ with the exception of end date which was extended to July 2007.

Eleven specialist databases were searched, including NHS Economic Evaluation Database (NHS EED), Health Economic Evaluations Database (HEED), Database of Abstracts of reviews of effects (DARE), Health Technology Assessment database and the National Research Register. Most of these databases were started in 1995 but some contained studies dating from 1993. They were searched without date restrictions to July 2007. In addition MEDLINE was searched from 1966-1994.

3.3.4 Checklist development

A checklist was used to assess the quality of the studies included in the review. The checklist is presented in Appendix B.⁷

Guidelines, checklists and criteria lists used to assess methodological quality were identified.¹⁰⁶ Chiou et al (2003) identified 19 lists after a systematic search of the English language literature since 1990.¹⁰⁷ Copies of these lists were used as the basis for the criteria for self care evaluations. There is some agreement across the 19 lists examined by Chiou et al (2003)¹⁰⁷ on the questions to be asked of economic evaluation studies. The general issues to be addressed in economic evaluations are largely the same across all types of interventions and the checklist presented in Appendix B reflects this. The check lists from Chiou et al (2003),¹⁰⁷ Drummond et al (1997),⁹⁹ NHS Centre for Reviews and Dissemination (2001),¹⁰⁸ and Forbes et al (2002)¹⁰⁹ were most influential in the checklist presented in Appendix B. In particular, a criterion was included in the checklist if it was judged to be either:

- a) important in assessing the quality of economic evaluations in general (for example the perspective of the study)
- b) particularly relevant to the evaluation of interventions to support self care (for instance the measurement of costs to include patients' out-of-pocket expenditure).

As with the Drummond checklist, there are several sub questions under some of the main questions. For these questions an assessment (subjective in nature) was made to assess whether the paper met the quality criteria. Such subjective assessments are a feature of most quality assessment checklists in the economic evaluation literature.

3.4 Results

The systematic literature search produced 2,570 abstracts. Of these 157 papers were deemed potentially includable in the review and the full paper was ordered. On examination of the full papers, 53 papers were identified as full economic

evaluations of self care support interventions and were therefore considered suitable for review. The studies identified as full economic evaluations and the paper selection method is illustrated in Figure 3.1 below:

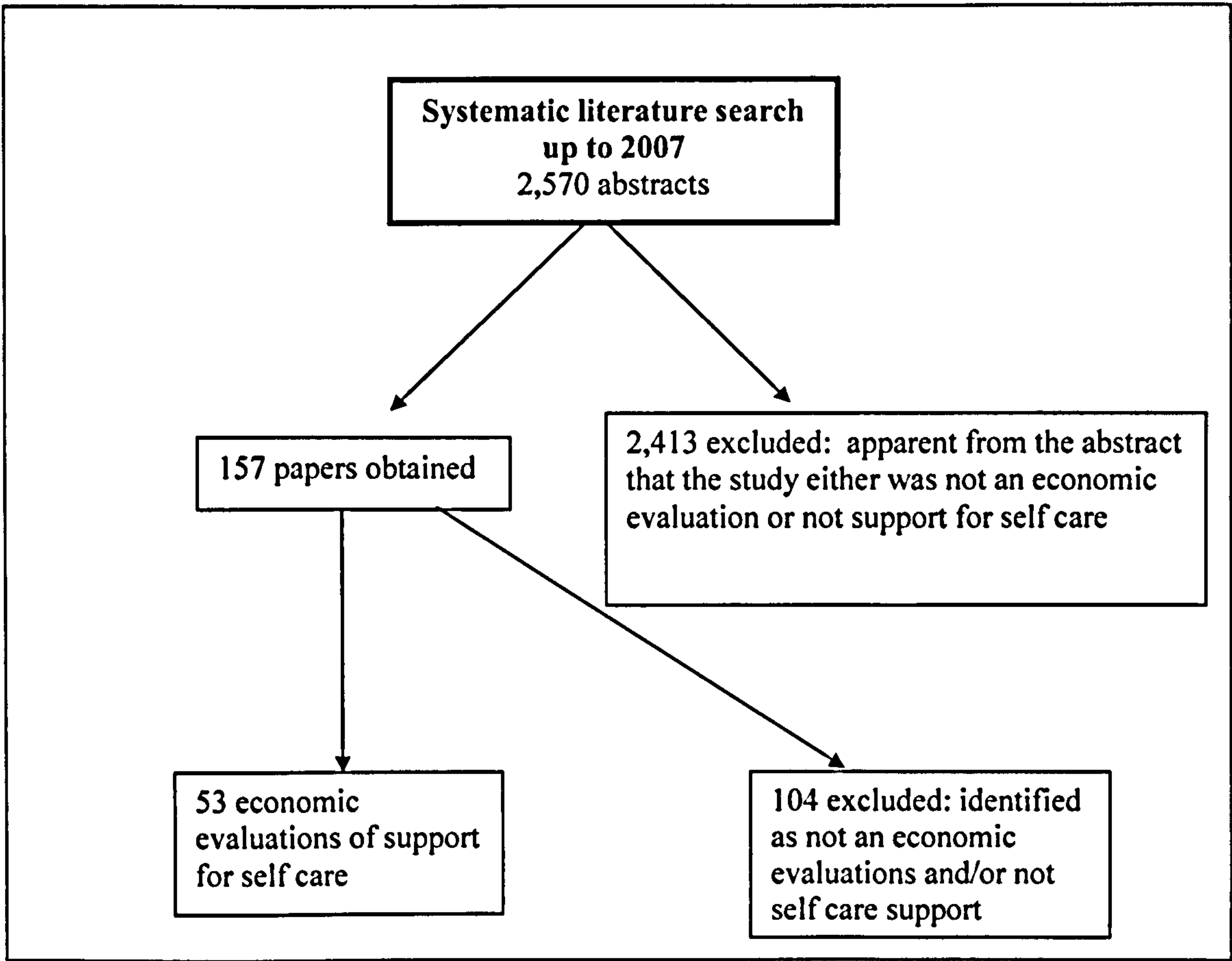


Figure 3.1 Selection of papers for assessment against criteria list

3.4.1 Study characteristics

Summary tables of the characteristics of full economic evaluations identified are provided below in Tables 3.1 and 3.2 including details of each study including the origin of the data (both country and date), the type of study (both economic and clinical), the geographical location and clinical setting. 22 of the 53 studies (42%) were based in the US, with eleven from the UK (21%), nine from Scandinavian countries (17%), and the remainder from other countries. All the studies were published after 1993 and the RCT (n=33, 62%) was the most common design of the clinical study.

Of the 53 studies, only 21 (40%) could be considered cost-effectiveness analyses or cost-utility analyses. The remainder were cost consequence analyses where the incremental costs and outcomes were not formally synthesised and/or compared,^{99, 110} rather a range of health (and/or other) outcomes are presented to decision makers.¹¹⁰

These interventions were applied across a range of (mainly chronic) conditions. Asthma (n=13), diabetes (n=8), arthritis (n=3) and heart disease (n=3) were the most commonly specified conditions, while “chronic disease” was evaluated in a further four studies.

Four studies did not report the follow-up period. Of the remainder, 12 months was the most common follow-up period (n=25), with 13 studies having a follow-up period (or time horizon for models) in excess of one year.

Table 3.1. Summary of economic evaluation papers (shaded rows provide a summary of studies identified since the earlier paper⁷ had been published)

Author	Pub date	Date of clinic al data	Origin	Condition	Setting for intervention	Clinical Outcome(s) measured	Comparators	E E type	Intervention
Albisser et al ¹¹¹	2001	NS	USA	Diabetes mellitus	Mixed model HMO	Blood glucose Body weight	1) Education 2) Education with self management training 3) Education with computer assisted self care	CCA	Education, self management training & computer-assisted self management on outcomes in diabetes disease management
Berg et al ¹¹²	2002	2000	USA	Diabetes	Community	Uptake of haemoglobin test Symptoms of hyperglycaemia BMI BP % on specific drugs	Historical control	CCA	Diabetes disease management
Cline et al ¹¹³	1998	1991 - 1993	Sweden	Heart disease	Hospital	1 year Survival rate drug treatments	1)Education and self management training 2) Usual care	CCA	Education and self management, plus easy access outpatient clinic.
Elston Lafata et al ¹¹⁴	2000	NS	USA	Heart	Hospital	Thromboembolic events Hemorrhagic events	1) Usual care 2) Clinic testing 3) Self testing	CCA	Anticoagulation clinics & patient self testing for patients on chronic Warfarin therapy.
Enght al ¹¹⁵	2001	NS	USA	Hip	Community	% in target therapeutic range DVTs Bleeding complications	1) Usual Care (prothrombin measured by nurse) 2) Self measurement	CCA	Self testing of prothrombin time after hip arthroplasty
Fitzmaurice et al ¹¹⁶	2002	NS	UK	CV	Primary care	% in target therapeutic range	1) Usual Care 2) Self management	CCA	Self care of anticoagulation vs primary care management

Table 3.1. Summary of economic evaluation papers

Gallefoss et al ¹¹⁷	2001	1994-95	Norway	Asthma	Outpatient dept	Lung function (FEV) QoL (SGRQ)	1) Education and GP Support 2) referred to GP	CEA	Self management programme.	education
Gallefoss et al ¹¹⁸	2002	1994-1995	Norway	COPD	Outpatient dept	NNE to make one independent of GP	1) Education and self management plan with follow-up by GP 2) follow-up from GP	CEA	Self management programme.	education
Ghosh et al ¹¹⁹	1998	1991-1994	India	Asthma	Hospital	Peak Expiratory Flow Rate (PEFR)	1) Regular care 2) Regular care plus self management training sessions	CCA	Self management asthmatics	training for
Glasgow et al ¹²⁰	1997	NS	USA	Diabetes	Primary Care	Dietary behaviour BMI Serum cholesterol Hyperglycaemic level	1) Usual Care 2) Brief intervention including goal setting and self help advise	CEA	Behavioural dietary intervention.	
Gray et al ¹²¹	2000	NS	UK	Type 2 diabetes	Hospital	Event Free year	1) Diet 2) Insulin and other drug control	CEA	Intensive blood glucose control policy vs conventional dietary control	
Groessl et al ¹²²	2000	NS	USA	Chronic illness	HMO	Cohesiveness Helplessness Knowledge Self-efficacy (Arthritis self-efficacy scale) Health Status (QWB) Intervention evaluation	1) Social support 2) Education 3) Combination of 1) and 2) 4) control group receiving periodic assessments	CCA	Social support and/or education.	

Table 3.1. Summary of economic evaluation papers

Humphreys ¹²³	2001	NS	USA	Subst abuse	Hospital	Abstinence Participation in self help groups	1) Cognitive behavioural interventions 2) Step based approaches	CCA	Comparison of professional encouragement of participation in 12 step self help groups vs cognitive behavioural therapy.
Jacobsen et al ¹²⁴	2002	NS	USA	Cancer	Hospital	QoL (SF36) Depression (CES-D) Anxiety (STAI-S)	1) Usual psychosocial care 2) professionally delivered self management 3) self administered stress management	CCA	Professionally administered stress management training or self administered stress management training.
Kauppinen et al ¹²⁵	1998	1991 - 1993	Finland	Asthma	Hospital	Lung function (FEV) QoL (15D and SGRQ)	1) Training including self management 2) Conventional education	CEA	Intensive vs conventional patient education & supervision for self management.
Kauppinen et al ¹²⁶	1999	1991 - 1993	Finland	Asthma	Hospital	Lung function (FEV) QoL (15D and SGRQ)	1) Training including self management 2) Conventional education	CCA	Intensive vs conventional patient education & supervision for self management.
Kauppinen et al ¹⁰⁵	2001	NS	Finland	Asthma	Hospital	Lung function (FEV) QoL (15D and SGRQ)	1) Training including self management 2) Conventional education	CEA	Intensive vs conventional patient education & supervision for self management.
Kruger et al ¹²⁷	1998	NS	USA	Arthritis	Community	Perceived pain	Model where one group gets Arthritis self help course compared to not getting it	CEA	Arthritis self help course as adjuvant to conventional therapy.
Lahdensuo et al ¹²⁸	1998	NS	Finland	Asthma	Community	Healthy days	1) Guided self management 2) traditional care	CEA	Guided self management

Table 3.1. Summary of economic evaluation papers

Lord et al ¹²⁹	1999	1995	UK	OA of knee	Primary care	SF36 General Health Questionnaire Arthritis Helplessness Index WOMAC	1) nurse led intervention for OA of knee 2) usual care	CEA	Nurse led education programme
Lorig et al ¹³¹	1993	1984	USA	Arthritis	Community	Health Status Pain Disability	1) Arthritis self management programme (ASMP) 2) Not getting this	CCA	Arthritis self management program
Lorig et al ¹³³	1999	NS	USA	Chronic disease	Community	Health behaviour Self-efficacy Health status	Historical control	CCA	Chronic disease self management program (CDSMP)
Lorig et al ¹³⁵	2001a	NS	USA	Chronic disease	Community	Health behaviour Health status	1) Chronic disease self management program (CDSMP) 2) not getting above	CCA	Chronic disease self management program (CDSMP)
Lorig et al ¹³⁹	2001b	1997	USA	Chronic disease	Hospital	Health behaviour Self-efficacy Health status	1) Chronic disease self management program (CDSMP) 2) not getting above	CCA	Chronic disease self management program (CDSMP).
Mattson Prince et al ¹³⁰	1995	1993	USA	Tetrapleg	Community	Health status RAND36, LSA-I, CHART, PASI, PIP) QoL	1) Agency provided personal assistance 2) self managed care	CCA	Self managed versus agency-provided personal assistance care.
Mazzuca et al ¹³¹	1999	1993	USA	OA of knee	Hospital	Pain Disability	1)self care education 2) attention control	CCA	Individualised self management education vs standard public education.

Table 3.1. Summary of economic evaluation papers

Meier ¹³² et al	2002	1997	USA	Type 2 diabetes	Hospital	Hyperglycaemic index Frequency of self managed blood glucose	Comparison under different guidelines	CCA	Modification of guidelines to reduce the number of self monitoring tests of blood glucose
-	-	1998							
Neri ¹³³ et al	1996	1993	Italy	Asthma	Hospital	Asthma attacks Medical examinations Admission days Lost working days	1) Educational program including asthma school 2) Reduced programme (booklet)	CEA	Teaching programs for asthma
-	-	1994							
Robinson ¹³⁴ et al	2001	NS	UK	Ulcer colitis	Hospital	Interval between relapse and treatment QoL (disease specific) Acceptability	1) TAU 2) consultation on request	CCA	Guided self management & patient-directed follow-up of ulcerative colitis
Rubin ¹³⁵ et al	1996	NS	USA	Urinary tract infection	Community	UTI avoided	2 alternative methods of OTC sale of oral antibiotics	CEA	Making oral antibiotics for urinary tract infection treatment available over the counter.
Salkeld ¹³⁶ et al	1997	1990	Australia	Heart disease	Primary care	Blood pressure BMI Cholesterol Smoking status QALYs	1) Routine care 2) GP Education plan 3) 2) and a program for risk factor behaviour change	CUA	General-practice based lifestyle change programs for people with risk factors for CV disease
-	-	1991							
Schermer ¹³⁷	2002	1996	Holland	Asthma	Health care	QALYs	1) Guided self management 2) Usual care	CUA	Guided self management of asthma in primary health care
-	-	1999							
Sinclair ¹³⁸ et al	1999	NS	UK	Smoking cessation	Community pharmacy	Smoking status	1) Routine provision 2) Train pharmacists to give quitting advise	CEA	Training of community pharmacists to deliver advice to customers on smoking cessation based on "state of change" model.

Table 3.1. Summary of economic evaluation papers

Starostina et al ¹³⁹	1994	NS	Russia	Type 1 diabetes	Hospital	No of daily insulin injections Daily insulin dose BMI Hyperglycaemic index Cases of severe hypoglycaemia and diabetic ketoacidosis. Knowledge test scores Hospital days	1) urine glucose self monitoring 2) blood glucose self monitoring	CCA	Intensive treatment and teaching programmes for type 1 (insulin dependent) diabetes mellitus comparing blood glucose to urine glucose self monitoring
Taborski et al ¹⁴⁰	1999	NS	Germany	Heart	Hospital	Complications	1) anti-coagulation self management 2) management by physician	CCA	Self managed anticoagulant therapy
Tschopp et al ¹⁴¹	2002	NS	Switzerl and	Asthma	Health care & community	Disease specific QoL	1) Self management intervention programme Historical control	CCA	Self management education booklet.
Volsko ¹⁴²	1998	NS	USA	Paed asthma	Hospital	Hospitalisations	Historical control	CCA	Education for self management
Von Korff et al ¹⁴³	1994	1989 - 1990	USA	Back pain	Primary care	Activity limitation	1) Group of GP practices who were consistent with self care 2) less towards self care	CCA	Practice style in managing back pain
Watson et al ¹⁴⁴	2002	2000	UK	Anti-fungals	Community	Appropriateness of management of thrush	Comparison of guidelines	2 CCA	Educational outreach for community pharmacists to promote evidenced-based practice

Table 3.1. Summary of economic evaluation papers

Castro et al ¹⁴⁵	2004	1996	USA	Asthma	Hospital	Asthma of Life	Usual Care	CCA	Nurse facilitated asthma education and support
Dunagan et al ¹⁴⁶	2005	1999	USA	Heart Failure	Community	SF12 MLHFQ	Usual Care	CCA	Nurse supported telephone based self care program
Gallefoss ¹⁴⁷	2004	1994/5	Norway	COPD	Hospital Outpatient	Patient satisfaction Need for reliever medication	Usual GP care	CCA	Group patient education with nurse and physiotherapist individual sessions
Jowett et al ¹⁴⁸	2006	NS	UK	Anticoagulation therapy	Primary Care	EQ-5D	Usual Care	CEA	Patient self management and training package
Keers et al ¹⁴⁹	2005	NS	Netherlands	Diabetes	Hospital (Outpatient)	Glycemic control (HbA1c) Diabetes related Distress	Usual care	CCA	Multidisciplinary education programme
Lindberg et al ¹⁵⁰	2002	NS	Norway	Asthma	Primary Care	Asthma related symptoms	Usual care	CCA	Nurse facilitated asthma education and support
Monninkhof et al ¹⁵¹	2004	NS	Netherlands	COPD	Secondary care	EQ-5D SGRQ	Usual care	CEA/ CUA	Self management guidelines and fitness programme
Mortimer & Kelly ¹⁵²	2006	NS	Aus	Diabetes	Community	Responder (defined as a well-managed diabetic)	Usual Care	CEA	Group based educational and social package
Richards et al ¹⁵³	2003	NS	UK	Mental health	Primary Care	EQ-5D, CORE	Usual GP care	CCA	Nurse facilitated self help programme with booklet
Richardson et al ¹⁵⁴	2006	NS	UK	IBD	Secondary Care	EQ-5D	Usual care	CUA	Whole system approach
Robinson et al ¹⁵⁵	2006	NS	UK	IBS	Primary Care	IBS QoL	Usual care	CCA	Self help guidebook and group therapy
Strong et al ¹⁵⁶	2006	1996	USA	Back Pain	Primary Care	Low impact back pain days	Usual Care	CEA	2 interventions based on lay led self management and psychologist led self management

Table 3.1. Summary of economic evaluation papers

Tagaya et al ¹⁵⁷	2005	2003	Japan	Asthma	Community	Peak Flow (PEF)	Usual Care	CEA	Asthma patient education program consisting mainly it appears of a booklet on symptoms, management of asthma
Waxman et al ¹⁵⁸	2003	1999 - 2000	UK	Foot care	Hospital Outpatient	Manchester foot disability score Patient Knowledge score Foot morbidity Patient enablement instrument	Usual Care	CCA	Self management program of education and training in foot care

Notes:

NS: Not stated

CCA: Cost consequences analysis. Costs and effectiveness (consequences) are presented separately ie are not synthesised/not presented as a ratio.

CEA: Cost-effectiveness analysis. Costs are expressed in monetary units and effectiveness is expressed in some single unit of effectiveness. When comparing two interventions the difference in cost and effectiveness between the two interventions is expressed as an incremental cost-effectiveness ratio (ICER), with the difference in cost in the numerator and the difference in effectiveness in the denominator

CUA: Cost-utility analysis. A form of CEA in which the units of effectiveness are quality-adjusted life-years (QALYs).

Table 3.2 Summary performance of economic evaluation papers against checklist

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q7	Q8	Q8	Q9	Q10	Q10	Q11	Q12	Q13	Q14	Q15	Q15	Q16	Q16	Q16	Q16
							a	b	a	b		a	b					a	b	a	b	c	
Albisser et al ¹¹¹	✓	✓	4	7	10	✓	X	✓	X	✓	✓	NA	NA	NA	X	X	X	NA	✓	✓	12	✓	
Berg et al ¹¹²	X	✓	3	7	10	✓	X	✓	X	NA	✓	NA	NA	NA	X	X	X	NA	✓	✓	13	✓	
Cline et al ¹¹³	✓	✓	3	5	10	✓	✓	X	X	X	X	NA	NA	NA	X	X	X	NA	✓	✓	12	✓	
Elston Lafata et al ¹¹⁴	✓	✓	1	8	10	✓	✓	✓	X	✓	✓	NA	NA	✓	✓	✓	NA	✓	✓	✓	14	✓	
Engh et al ¹¹⁵	✓	✓	3	6	10	✓	✓	✓	X	X	✓	NA	NA	NA	X	X	X	✓	✓	✓	13	✓	
Fitzmaurice et al ¹¹⁶	✓	✓	3	5	10	X	X	✓	X	X	✓	NA	NA	NA	X	X	X	NA	✓	✓	13	✓	
Gallefoss et al ¹¹⁷	✓	✓	1	5	9	✓	✓	✓	✓	✓	✓	NA	NA	NA	✓	✓	✓	NA	✓	✓	12	✓	
Gallefoss et al ¹¹⁸	✓	✓	1	5	9	✓	✓	✓	X	✓	✓	NA	NA	NA	✓	✓	✓	NA	✓	✓	12	✓	
Ghosh et al ¹¹⁹	✓	✓	2	5	10	✓	✓	✓	✓	✓	✓	NA	NA	NA	X	X	X	NA	✓	✓	12	✓	
Glasgow et al ¹²⁰	✓	✓	3	5	9	✓	X	✓	✓	✓	✓	NA	NA	NA	✓	✓	X	NA	✓	✓	12	✓	
Gray et al ¹²¹	✓	✓	3	5	9	✓	✓	✓	✓	✓	✓	NA	NA	✓	✓	✓	✓	NA	✓	✓	12	✓	
Groessl et al ¹²²	✓	✓	3	5	10	✓	✓	✓	✓	✓	✓	NA	NA	✓	✓	✓	✓	NA	✓	✓	13	✓	
Humphreys ¹²³	✓	✓	1	6	10	X	X	X	X	X	X	NA	NA	NA	X	X	X	NA	✓	✓	12	✓	
Jacobsen et al ¹²⁴	✓	✓	3	5	10	✓	X	✓	✓	✓	✓	NA	NA	NA	✓	✓	X	NA	✓	✓	12	✓	
Kauppinen et al ¹²⁵	✓	✓	2	5	9	✓	✓	✓	✓	✓	X	NA	NA	NA	✓	✓	X	NA	✓	✓	12	✓	
Kauppinen et al ¹²⁶	✓	✓	3	5	9	✓	✓	✓	✓	✓	✓	NA	NA	NA	✓	✓	X	NA	✓	✓	12	✓	
Kauppinen et al ¹²⁶	✓	✓	3	5	9	✓	X	✓	✓	✓	✓	NA	NA	✓	✓	✓	X	NA	✓	✓	12	✓	
Kauppinen et al ¹²⁶	X	✓	4	5	10	X	X	✓	✓	✓	X	NA	NA	✓	X	✓	X	NA	✓	✓	12	✓	
Kruger et al ¹²⁷	✓	✓	1	8	9	✓	✓	X	✓	✓	X	NA	NA	✓	✓	✓	NA	✓	✓	✓	14	✓	
Lahdensuo et al ¹²⁸	✓	✓	3	5	9	✓	X	X	✓	✓	X	NA	NA	NA	✓	X	X	NA	X	X	14	X	
Lord et al ¹²⁹	✓	✓	1	5	9	✓	✓	✓	✓	✓	✓	NA	NA	NA	✓	✓	X	NA	✓	✓	12	✓	
Lorig et al ¹²¹	✓	✓	4	7	10	X	X	✓	X	✓	✓	NA	NA	✓	X	✓	X	NA	✓	✓	12	✓	
Lorig et al ¹²³	✓	✓	3	5	10	✓	✓	✓	✓	✓	X	NA	NA	✓	✓	X	X	NA	✓	✓	12	✓	
Lorig et al ¹²³	✓	✓	3	5	10	✓	✓	✓	✓	✓	X	NA	NA	NA	✓	X	X	NA	✓	✓	12	✓	
Lorig et al ¹²³	✓	✓	3	5	9	✓	✓	✓	✓	✓	X	NA	NA	NA	✓	X	X	NA	✓	✓	12	✓	
Lorig et al ¹²³	✓	X	3	7	9	✓	X	X	✓	NA	✓	NA	NA	NA	✓	X	X	NA	✓	✓	13	✓	
Long et al ¹²⁹	✓	✓	3	6	10	✓	✓	✓	X	X	✓	NA	NA	NA	X	✓	X	NA	✓	✓	12	✓	
Mattison Prince et al ¹³⁰	✓	✓	3	6	10	✓	✓	✓	✓	✓	✓	NA	NA	NA	X	✓	X	NA	✓	✓	12	✓	
Mazzuca et al ¹³¹	✓	✓	2	6	10	✓	✓	✓	✓	✓	✓	NA	NA	NA	X	✓	X	NA	✓	✓	12	✓	
Meier et al ¹³²	✓	✓	3	7	10	✓	✓	✓	✓	X	✓	NA	NA	NA	X	X	X	NA	✓	✓	12	✓	
Neri et al ¹³³	✓	✓	4	5	9	✓	X	✓	✓	✓	✓	NA	NA	NA	✓	✓	X	NA	✓	✓	12	✓	
Robinson et al ¹³⁴	✓	✓	4	5	10	✓	X	X	X	X	X	NA	NA	NA	X	X	X	NA	X	X	12	X	

✓ = Yes , X = No, NA = Not applicable, NS = Not stated		
<i>Q1 Study clarity</i>	<i>Q5 Economic study design</i>	<i>Q11 Discounting</i>
<i>Q2 Comprehensive description of competing alternatives</i>	9 = Cost-effectiveness Analysis (CEA) (31%)	<i>Q12 Incremental analysis</i>
<i>Q3 Perspective</i>	10 = Cost Consequence Analysis (CCA) (58%)	<i>Q13 Allowance for uncertainty</i>
1 = Societal (30%)	11 = Cost-utility Analysis (CUA) (11%)	<i>Q14 Missing data handled appropriately</i>
2 = Health care system & patient (6%)	<i>Q6 Design adequacy given study type</i>	<i>Q15a Economic model</i>
3 = Health care system (54%)	<i>Q7a Relevant costs identified</i>	<i>Q15b Appropriateness of economic model</i>
4 = Not clear (11%)	<i>Q7b Relevant consequences identified</i>	<i>Q16a Type of funder</i>
<i>Q4 Study design</i>	<i>Q8a Costs measured accurately</i>	12 = Public/voluntary sector (73%)
5 = Randomised Control Trial (RCT) (62%)	<i>Q8b Consequences measured adequately</i>	13 = Private sector (13%)
6 = Case Control Trial (CCT) (11%)	<i>Q9 Statistical analysis appropriateness given the design</i>	14 = Do not state (13%)
7 = Before and after (21%)	<i>Q10a Sub-group analysis</i>	<i>Q16b Generalisability</i>
8 = Decision model (6%)	<i>Q10b Sub-groups pre-specified</i>	<i>Q16c Presentation and discussion of key results</i>

3.4.2 Study conclusions and quality

Most of the studies concluded that the self care support intervention was either dominant (in that costs were reduced and patient outcomes improved) or cost-effective (the additional benefit was judged to be worth the extra cost) (n=41, 77%). Only four studies claimed that the intervention was not cost-effective, while the remainder (n=8) were inconclusive. While this may appear to be in favour of evidence to support these interventions, most studies had significant flaws, which have been previously discussed.⁷ The most common flaws were:

- a) Poor costing methodology. Only 30% of studies (n=16) had a societal perspective. Others defined costs narrowly, for example ignoring patient expenditure. Several also did not present unit cost data so that replication of results was not possible.
- b) Inappropriate comparison group. Several studies (n=11) used a before and after design. There may be many reasons for changes over time in such studies, for example, regression to the mean may be an important factor where costs are high in one period and lower in the next.
- c) Inadequate handling of uncertainty. Many studies did not conduct sensitivity analysis (n=20) nor presented confidence intervals around mean estimates of cost or effect. Without these analyses/estimates, it is not possible to estimate the probability that an intervention is cost-effective.
- d) Missing data were either ignored or dealt with inappropriately. Few studies (n=8) reviewed handled missing data using recommended techniques (such as multiple imputation). It has been shown that results can be sensitive to imputation of missing data and to the choice of imputation method employed.¹⁵⁹
- e) Short period of follow-up. The majority of studies (n = 38) had a follow-up (or time horizon) of one year or less. The majority of clinical trials have relatively short follow-up periods. This “truncation of the time horizon” has been

discussed in more detail in section 2.3.2, but can lead to erroneous estimates of the treatment effect.

- f) Choice of outcome measure. Because studies often aim to demonstrate improvements not only in patients' health but also in empowerment, autonomy and control of health affecting decisions, evaluations tend to include a wider range of outcome. While this is not necessarily a weakness of the study, it is a limitation to the generalisability of the study and its usefulness for decision making (as these outcomes are not used in different conditions and are seldom valued).

Eleven UK based studies were identified. Of these, five studies *claimed* that the intervention to support self care was cost-effective, in that either the intervention was dominant (costs reduced and outcomes improved) or the incremental cost-effectiveness ratio (ICER) lay below a specified threshold value.^{134, 153-155, 158} In one case this is a questionable claim,¹⁵⁸ as the study reported increased costs and improvement in outcomes, but no valuation of outcomes was performed. Three studies reported that support for self care was not (or was unlikely to be) cost-effective as the ICER exceeded a specified threshold.^{116, 129, 148} Three studies were inconclusive as costs and outcomes were both increased and no valuation of outcomes was performed to assess whether the benefits might outweigh the costs.^{121, 138, 144}

The UK based studies were, in general, well conducted with two presenting uncertainty in the form of cost-effectiveness acceptability curves and dealing with missing data in an appropriate manner.^{148, 154} Clearly, the proportion of UK studies claiming the intervention was cost-effective (5/11, 45%) was lower than that of the whole sample (41/53, 77%). Similarly, the proportion of studies claiming that the intervention was not cost-effective was considerably higher in the UK sample (27% vs 8%), though sample sizes are too small to make any peremptory statements regarding cost-effectiveness of these interventions in the UK. Whether it was the better quality of the UK studies, or the geographical location that influenced the lack of cost-effectiveness is unclear given the existing evidence.

The evaluations of interventions based on the Chronic Disease Self Management Program or the Arthritis Disease Management Programme were not generally well conducted. Four economic evaluations of this intervention (on which the UK Expert Patients Programme is based) were identified.^{71, 75, 79, 80} These studies were all conducted in the US in patients with a range of chronic conditions. These studies were poorly conducted in that missing data were largely ignored rendering the statistical analysis largely questionable and in general had little allowance for uncertainty. Further *clinical* evaluations of these interventions are reviewed in Chapter 6, but evidence of the cost-effectiveness of these interventions from the existing published literature must be considered to be unreliable.

3.5 Discussion

Interventions to support patient self care are very diverse and take place in many types of setting, so drawing general conclusions about the cost-effectiveness of these interventions from the existing literature is problematic. The pooling of results of these studies is inappropriate due to the heterogeneous nature of the studies, their patient populations, settings and choice of outcome measure.

There were clear differences in the results from studies in different countries. Three of the four studies that showed that these interventions were not cost-effective were UK based, while the proportion claiming that these interventions were cost-effective was noticeably lower than for the whole sample.

It is not possible to ascertain whether the contrasting conclusions generated from the different geographical locations of these studies is due to the different quality of evaluations carried out in different countries or a real difference in the relative cost-effectiveness of these interventions in different countries.

The cost-effectiveness of interventions may also differ between conditions. For example, there were four evaluations of chronic illness, all claiming that the interventions were cost-effective.^{75, 79, 80} However, for osteoarthritis of the knee, there were two studies,^{129, 131} one of which showed equivocal results¹³¹ the other demonstrating that the self care intervention was not cost-effective.¹²⁹ Similarly, the

type of intervention is likely to impact on effectiveness and cost-effectiveness. For example, groups providing training in self care may be very different in their effectiveness and cost-effectiveness to pharmacists providing information for individuals.

Resource and unit cost data may not be readily transferable between systems. It is also likely that the outcome effects of these interventions may be culturally dependent. Some interventions concentrate on a very narrow patient population while others consider all chronic conditions. It is unsurprising that the conclusions of studies based on these diverse populations are different. There may also be more control group contamination in some studies than in others (for example where the intervention is based on a published guidebook, it may be possible for control group patients to access that information). These issues, particularly the transferability of treatment effects on costs and outcomes, are discussed in greater detail in Chapter 6.

Many of the studies reviewed were of poor quality, most having one or more major drawbacks. Many of these studies would not be of the quality required by NICE for use in technology appraisal.¹² This lack of quality limits their usefulness for UK policy. A previous review of evaluations of self management interventions was also critical of the methodology and of the lack of consideration of the cost-effectiveness of self management interventions.¹⁶⁰ Bower et al noted that there were no data on long-term clinical or cost-effectiveness for these interventions and “available evidence is limited in quantity and quality and more rigorous trials are required to provide more reliable estimates of the clinical and cost-effectiveness of these treatments”.¹⁶¹

3.6 Conclusion

Chapter 2 identified the various paradigms within economic evaluation. The majority of economic evaluations reviewed in this chapter were based in the extra-welfarist paradigm (either implicitly or explicitly) with the maximisation of patients' health outcomes (rather than changes in individuals utility), subject to an exogenously given budget constraint. None of the economic evaluations identified sought to broaden the perspective of the study beyond this approach. In addition, the majority

of evaluations were based on the results of a single trial/study, with little attempt to synthesise data. The majority of evaluations were based in the frequentist paradigm with few studies providing estimates of, for example, the probability of the intervention being cost-effective at various values of the outcome measure chosen, though there were exceptions to this.^{148, 162}

There was no evidence of the identified studies attempting to synthesise all relevant evidence, to use an outcome measure that is useful for decision making purposes or to place a “value “ on the outcomes measured. The usefulness of the studies described above for informing the debate over whether interventions to support self care are cost-effective in the UK setting is limited. To be helpful for decision makers to choose between interventions across a variety of conditions, results of evaluations should use a generic outcomes measure.^{12, 163} The QALY is an example of generic measure of health related quality of life encompassing both the quality and quantity of life. Chapter 4 presents an economic evaluation using the QALY as an outcome measure. This cost-effectiveness study is based on the results of a single randomised controlled trial conducted in the extra-welfarist tradition with maximisation of health (as measured by the QALY) as the outcome measure of interest. The chapter presents an evaluation of the Expert Patients Programme (EPP) that is now widely available in the NHS and is intended to be available to 100,000 individuals with chronic conditions by the year 2012.¹⁶⁴

Chapter 4. The Expert Patients Programme Cost-Effectiveness Evaluation.

4.1 Introduction

The previous chapter concluded that previous economic evaluations of self care support interventions were of limited use for decision makers in the UK, partly due to their poor quality and partly due to the use of outcome measures which are difficult to value. The decision maker's task of efficiently allocating resources is aided by the use of a generic outcome measure which has some commonly accepted value. This chapter describes the single trial based cost-effectiveness analysis undertaken alongside the randomised controlled trial of the Expert Patients Programme (EPP). The outcome measure for the economic analysis was a generic instrument (the QALY) which enables comparisons to be made between interventions in different conditions. As such, the results of an economic evaluation of EPP are directly relevant to the decision problem specified earlier and are therefore of interest to decision makers concerned with the cost-effectiveness of interventions to support self care.

The chapter employs an extra-welfarist perspective and utilises Bayesian methodology, as described in Chapter 2, to present an analysis that is intended to be useful for decision makers. The review presented in Chapter 3 identified several drawbacks from the existing literature which the analysis presented in this chapter avoids. Notably, previous economic evaluations of the EPP have tended to handle uncertainty and missing data inappropriately,^{71,73,75,79} as well as use an outcome that does not facilitate comparison between cost-effectiveness in differing disease areas.

The objective of this chapter is to present the aims, design, conduct and analysis of the cost-effectiveness of the EPP intervention based on a single trial based evaluation. The uncertainty associated with this estimate is presented by estimating the probability that the EPP is cost-effective over a range of values of decision makers' threshold value of an additional quality-adjusted life-year. The contents of this chapter form the basis of a publication in press with the Journal of Epidemiology

and Community Health.¹⁶² The cost-effectiveness results presented in this chapter were also included in the clinical paper previously published in the same journal.⁶⁶

4.2 Background to the EPP

The NHS Plan identified self care as an important factor in providing a “patient centred health service”.⁸ More recently a Department of Health report has claimed that supporting self care can “improve health outcomes [and], improve patient satisfaction”,³ while the roll-out of the Expert Patients Programme (EPP) “could create a generation of patients empowered to take action to improve their health”.⁹ ¹⁰ The ‘Expert Patients Programme’ (EPP) aims to provide self care support to any individual with a chronic condition in England. The EPP is based on a generic programme developed in the US, the chronic disease self management programme (CDSMP). This lay led group intervention is designed to enable participants to develop appropriate self care skills.⁶⁶

As asserted in Chapter 1, in a budget constrained system such as the National Health Service (NHS), it is important for these interventions to promote patients’ self care or self management to be cost-effective as well as clinically effective. To establish whether an intervention is cost-effective in a budget constrained system requires that the outcome measures used are generic.

4.3 Methods

The EPP is an intervention designed to support patients’ self care that has been “rolled out” across England and Wales. This analysis is conducted in the extra-welfarist tradition in that there is an assumed exogenously defined budget constraint and health as measured by the QALY is the maximand. This enables comparison across conditions which were not possible with many of the trials reported in the previous chapter.

This cost-effectiveness analysis takes a Bayesian perspective in that the parameters are considered to have probability distributions and that it is not a requirement to

have a set of random experiments (a sample space) from which to assess the relative frequency of events. In this paradigm it is possible, and appropriate, to compute the probability of an event being effective and/or cost-effective. It was argued in an earlier chapter that this is a more useful technique for presenting the results to decision makers.

Briefly, the EPP evaluation was a two-arm trial comparing the clinical and cost-effectiveness of a lay led self care support programme. The comparator was a waiting list control. It is acknowledged that some of the weaknesses identified in the literature review in Chapter 3 have been replicated in this cost-effectiveness study. In particular, the study has a relatively short follow-up period and the control group were a waiting list control rather than an active intervention. These were shortcomings that were unavoidable in this practical study. Nevertheless, the design, conduct and analysis of this study conform with accepted methodologies to a greater degree than did the previous studies. This is particularly the case in the treatment of missing data, the characterisation of uncertainty, the costing methodology and the choice of outcome measure.

The EPP evaluation can be described as a “pragmatic” trial or “primary cost-effectiveness trial” as it was based in a real world setting and considered cost-effectiveness rather than cost-efficacy.^{66, 165}

4.3.1 Participants

The intervention was designed to enhance the self-efficacy of patients with a self defined long-term condition. Any individual with a self defined long-term condition could participate in the trial. Participants in the trial self reported their main condition on the baseline questionnaire; other co-morbidity was also recorded but analysis was carried out on the main condition. There were no specific inclusion or exclusion criteria, with recruitment carried out in all 28 Strategic Health Authorities in England. Recruitment was community based and used a variety of methods including posters in GP surgeries and media advertisements. Thus, patients were not recruited in clusters (such as GP practice), but individually as they put themselves forward for the EPP program. Characteristics of the two groups were presented in the clinical paper,⁶⁶ and are reproduced below in Table 4.1. There was little difference between

the groups in types of condition.

The most common conditions were musculo-skeletal and endocrinal and were closely matched in the two groups. In addition, the groups were comparable in terms of their age, gender, ethnicity and other socio-demographic characteristics.

4.3.2 Intervention

The intervention consisted of six 2.5 hour group sessions, with between 8 and 12 individuals per group, and was held weekly. Topics within the sessions included relaxation, diet, exercise, fatigue, breaking the 'symptom cycle', managing pain and medication, and communication. Groups were led by two lay trainers or volunteer tutors. The lay leaders were people with lived experience of long-term conditions. They were trained and subject to a quality control process.¹⁶⁶ The intervention focussed on increasing participants' self-efficacy through problem solving and goal setting. Patients in the waiting list control could access the intervention after six months. While on the waiting list control, participants received treatment as usual and were advised to continue to manage their condition as they usually would.

4.3.3 Outcome measures

The outcomes used in the clinical trial reflected the objectives of the intervention. The mechanisms involved in the evaluation of complex interventions can be intricate.¹⁶⁷ As described in Chapter 1, the EPP is based on the Chronic Disease Self Management Programme that emphasises the importance of self-efficacy in the management of chronic conditions. Health psychology models of the CDSMP suggest a theoretical model where change in self-efficacy cognitions (patients' confidence in managing their condition) acts as a causal mechanism towards health status and utilisation.¹⁶⁸

Individuals' level of self-efficacy was included as a measure of outcome. Self-efficacy is measured on a 1 to 10 scale varying from "not at all confident" to "totally confident". In addition, the cost-effectiveness analysis used the QALY (described below) as a measure of outcome.

Table 4.1 Baseline demographic and health characteristics

<i>Characteristic</i>	<i>Intervention n = 313</i>		<i>Control n = 316</i>	
Age (SD)	55.5	(13.6)	55.3	(13.6)
Gender: Female	219	(70.0%)	220	(69.6%)
Ethnicity: White	298	(95.2%)	299	(94.6%)
Marital status: Lives alone	82	(26.2%)	93	(29.4%)
Lives with spouse/partner	188	(60.1%)	190	(60.1%)
Educational qualifications: None	77	(24.6%)	61	(19.3%)
Degree	51	(16.3%)	53	(16.8%)
Accommodation: Owner-occupied	214	(68.4%)	214	(67.7%)
Work situation: In paid work	58	(18.5%)	66	(20.9%)
Unable to work due to condition	111	(35.5%)	106	(33.5%)
Retired	110	(35.1%)	111	(35.1%)
Self reported main long-term health condition ¹ :				
Musculoskeletal	106	(33.9%)	107	(33.9%)
Endocrine	37	(11.8%)	37	(11.7%)
Circulatory	20	(6.4%)	24	(7.6%)
Myalgic encephalitis /chronic fatigue	22	(7.0%)	25	(7.9%)
Respiratory	23	(7.4%)	17	(5.4%)
Mental health	19	(6.1%)	19	(6.0%)
Neurological	20	(6.4%)	18	(5.7%)
Other	66	(21.1%)	69	(21.8%)
Self reported general health:				
Very good/excellent	32	(10.2%)	34	(10.8%)
Good	90	(28.8%)	92	(29.1%)
Fair	122	(39.0%)	111	(35.1%)
Poor	69	(22.0%)	79	(25.0%)
Self reported baseline health characteristics:				
Self-efficacy (SD)	45.9	(21.5)	47.7	(22.3)
Energy (SD)	32.6	(19.5)	33.3	(20.1)
Service utilisation (SD)	8.6	(7.3)	9.1	(8.1)
Classification of Primary Care Trust locality ² :				
Predominantly rural	94	(30.0%)	93	(29.4%)
Some rural and mixed	114	(36.4%)	118	(37.3%)
Major and large urban	105	(33.6%)	105	(33.2%)
Seasonal change between recruitment and follow-up ³ :				
Follow-up at a 'worse' month in cycle than recruitment	21	(6.7%)	18	(5.7%)
Follow-up at a 'better' month in cycle than recruitment	31	(9.9%)	45	(14.2%)
No difference	28	(9.0%)	24	(7.6%)
No seasonal pattern	233	(74.4%)	229	(72.5%)
Time between recruitment and follow-up: (mean days, SD)	219	(40.6)	209	(37.1)

¹ Post-hoc classification. The classification used in minimisation was: musculoskeletal, diabetes, heart disease, other.

² Definition taken from Department of Environment, Food and Rural Affairs classification of Primary Care Trusts in England (September 2005).

³ Based on informant ratings of each calendar month, collected at baseline.

Data for all outcome measures were collected at baseline, 6 months and 12 months. Postal questionnaires were used for all follow-ups, non-responders were subsequently contacted with telephone reminders.

At the 12 month follow-up the control group had been offered treatment and therefore the analysis below considers only the six month data.

4.3.4 Perspective

Patients were the unit of randomisation and the unit of analysis. The trial used a 1:1 randomisation ratio. The analysis takes a societal perspective (including the costs to patients) with effects assessed in terms of health gains, measured in terms of quality-adjusted life-years (QALYs). An additional analysis considers the perspective of government agencies (health and social care).

All costs and outcomes fell within a six month period and therefore discounting was not appropriate.

4.3.5 Resource Use and imputation

Traditionally, economic evaluations do not collect data on resource use before the intervention is delivered. Thus, the mean resource use in each group is measured without considering whether individuals in the group were substantial users of resources or not. However, more recently, the importance of adjusting for baseline imbalance has been recognised.⁴⁴ In this trial previous resource use data were collected and therefore could be used as predictors of final resource use. In the EPP evaluation, patients were asked to estimate their use of resources over the previous six months (i.e before randomisation). For each item of resource use, this pre-baseline value was used as a covariate to adjust post-randomisation resource use.

Therefore, two analyses of resource use are presented. Firstly, the analysis is presented in the usual manner without adjustment for previous resource use. Subsequently an analysis where resource use is adjusted for previous resource use is presented.

Analysis of resource use data using imputed values is problematic as pre-randomisation resource use was used as an explanatory variable in the imputation of missing values and therefore there would be an element of double counting. In addition, where multiple imputation techniques are used (see sensitivity analysis in

section 4.8), several (usually five) datasets would need to be created for *each* resource use, in this case generating around 100 datasets, which are unwieldy for analysis. Therefore, for the resource use estimates presented in Table 4.3, the figures are based on the complete case analysis.

4.4 Sources of data

4.4.1 Resource use

Resource use data were collected alongside clinical data during the EPP trial. Patient questionnaires (see Appendix G) were administered at baseline and six month follow-up. Patients were asked for details of visits to primary health care practitioners, secondary care appointments and hospital stays as well as community based support and individuals' out-of pocket expenditure. The medication that each patient received was also recorded.

4.4.2 Unit costs

Inpatient cost per day and outpatient cost per visit for attendances were both based on national estimates.¹⁶⁹ Estimates were inflated to a 2003/2004 price base using the Health Service Cost index. These figures are national estimates and are therefore appropriate for an evaluation conducted across all the SHAs in England.

The cost of a GP visit (both home and surgery) and the cost of a practice nurse visit (home and surgery) were derived from Curtis and Netten estimates.¹⁷⁰ The unit cost estimate includes cost of training as well as direct care support staff and is inflated to a 2003/2004 price base.

The unit cost of each medication was estimated from the British National Formulary.¹⁷¹

The cost of the intervention used in the trial was estimated from data on the overall costs of the EPP and the throughput over the period of the trial (the period April 2003-March 2005).¹⁷² These cost estimates include the cost of managing the programme, costs of training and delivery of the programme as well as the cost of providing facilities.

4.4.3 Health States and their value

The EQ-5D instrument was used to measure patients' health states and to ascribe those states values. EQ-5D measures patient health status across five dimensions (mobility, self care, usual activities, pain/discomfort and anxiety/depression) with three possible responses (no problems, moderate problems or severe problems) for each dimension.¹⁷³ This locates each participant into one of 245 mutually exclusive health states (with the addition of death and unconscious), each of which has previously been valued on the zero (equivalent to dead) to one (equivalent to full health) valuation scale based on interviews with a sample of 3,395 members of the UK public.¹⁷⁴ The questionnaire is presented in Figure 4.1 below.

QALYs were calculated by plotting the EQ-5D utility score at baseline and at six month follow-up and calculating the area under the curve. This is the appropriate method as it reflects the fact that the QALY is a product of both time and utility.¹⁷⁵ As is common in the health economics literature, it was assumed in the absence of evidence to the contrary, that changes in utility score over time followed a linear path,⁴³ that is there were no higher order terms (squared, cubed etc) included in the change in EQ-5D scores over time. This generates a QALY gained for each patient over the six month time period of the study. An illustration of the QALY calculation for the control group in this study is presented below. More detailed examination of QALY calculation can be found elsewhere.⁴³

The two trial groups (intervention and control) were then compared over the six month period to generate the estimate of mean differential QALYs between the groups. These estimates were then adjusted for baseline EQ-5D as recommended by Manca et al.⁴⁴ Gender and age were also included as covariates in order to adjust for any differences at baseline (whether statistically significant or not).

☐ Tick one box for each group of statements.

Mobility

- I have no problems in walking about
- I have some problems in walking about
- I am confined to bed

☐
☐
☐

Self-Care

- I have no problems with self-care
- I have some problems washing or dressing myself
- I am unable to wash or dress myself

☐
☐
☐

Usual Activities

- I have no problems with performing my usual activities
(e.g. work, study, housework, family or leisure activities)
- I have some problems with performing my usual activities
- I am unable to perform my usual activities

☐
☐
☐

Pain/Discomfort

- I have no pain or discomfort
- I have moderate pain or discomfort
- I have extreme pain or discomfort

☐
☐
☐

Anxiety/Depression

- I am not anxious or depressed
- I am moderately anxious or depressed
- I am extremely anxious or depressed

☐
☐
☐

Compared with my general level of health over the past 12 months, my health state today is:

- Better
- Much the same
- Worse

☐
☐
☐

Figure 4.1 The EQ-5D instrument

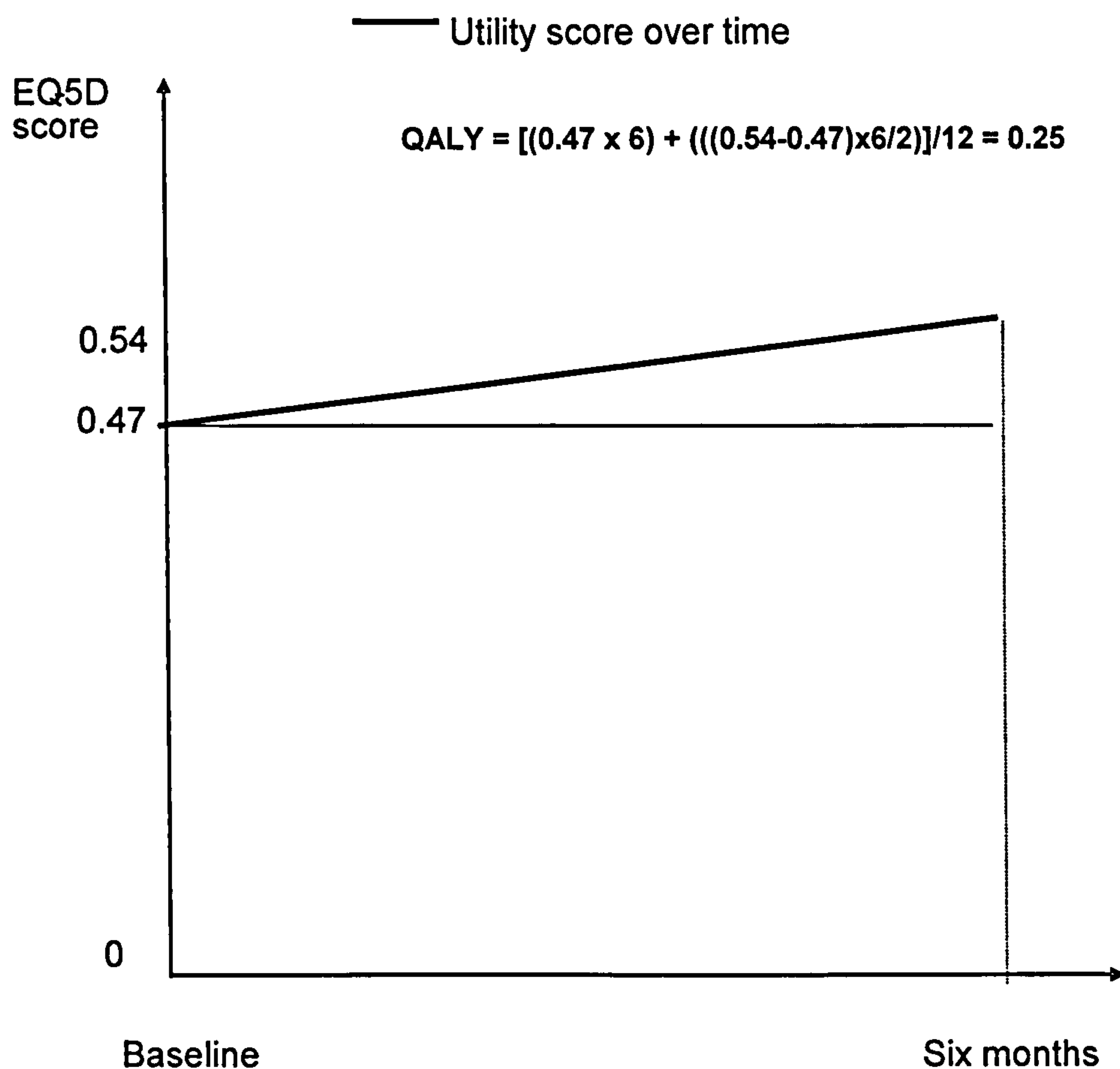


Figure 4.2 Calculation of a QALY

4.5 Methods of analysis

4.5.1 Missing data and imputation

Where resource use or EQ-5D data were missing, the analyst needs to be aware of the pattern of missing data. There is no formal test to verify the assumption that data are missing at random (MAR), and this assumption is often chosen as a starting point when data are missing.¹⁷⁶ This approach was employed in this analysis.

The best subset regression method was employed to impute missing data using Stata 8. The impute function in Stata conducts an efficient missing value

regression by sorting the missing data into patterns. This method determines which independent variables should be included in a regression model by examining all of the models created from all possible combinations of independent variables and uses R^2 to check for the best model.

Explanatory variables included in all imputations were treatment group, age and gender. For missing EQ-5D scores, the baseline score on that dimension was also employed as a covariate. For missing resource use data, the baseline value for that resource was included as a covariate.

Multiple imputation, where several datasets are created with different values generated for each missing value in each dataset, was employed as a sensitivity analysis.

4.5.2 Incremental cost-effectiveness ratios (ICERs) and net monetary benefits

The ICER is calculated from the mean difference in cost and effect between the two treatment options. The ICER is presented in this analysis where appropriate.

However, because of problems interpreting ICER statistics that cover more than one quadrant of the cost-effectiveness plane, this analysis also uses the net benefit approach to cost-effectiveness analysis described below.^{88, 90, 177}

Using the patient level estimates of costs and effects, it is possible to determine the joint density of costs and effects by re-sampling (the non-parametric bootstrap).¹⁷⁸

In this instance, replicated samples were made by drawing from the original sample (with replacement) of 629 patients' costs and effects (313 EPP and 316 control).

This procedure is repeated 1,000 times generating 1,000 distributions of costs and effects for each group. For each of these re-samples, the range of values of the cost-effectiveness threshold value of a QALY (henceforth represented by λ) was applied to the mean effects (reflecting the range of values that could be placed on the measure of effect, in this case the QALY), thereby generating the net benefit figure (for each value of λ). The cost-effectiveness threshold value of a QALY is a matter of some debate,⁶¹ with commonly quoted values of £20,000 to £30,000 per QALY.¹² Cost-effectiveness analyses commonly portray ranges of values of £0 to £50,000 for this threshold.

Thus, the value of net monetary benefit is dependent on the threshold value (λ) of an additional QALY (as below).

$$\text{Net Monetary Benefit (NMB)} = (\lambda * \text{QALYs}) - \text{Cost}$$

If the Incremental Net (monetary) Benefit (INB) (i.e. the NMB in the EPP group minus the NMB in the control group) is greater than zero for that value of λ , the treatment is judged to be cost-effective. The proportion of the samples of the INB estimates that are judged to be cost-effective over the range of values of threshold values of a QALY is then used to generate the cost-effectiveness acceptability curve (CEAC).^{88, 90}

Thus the uncertainty surrounding the NMB statistic can be used to identify the probability that a strategy is cost-effective using the cost-effectiveness acceptability curve (CEAC). The CEAC is a graphical representation of the probability of an intervention being cost-effective over a range of monetary values for threshold value of a QALY. The probability of an intervention being cost-effective differs according to the valuation the decision maker places on a QALY.

As an example, imagine we have an intervention that improves QALYs but also costs more. It may be that if we place a value of £10,000 on a QALY, 75% of the replications result in a positive NMB. However, if we value gains in QALYs more than this, say at £30,000 per QALY, then 95% of replications result in a positive NMB. The CEAC portrays these estimates in graphical format.

For this analysis the threshold value of a QALY was varied from zero to £50,000. The value zero is equivalent to a comparison of the groups in terms of total costs, as outcomes are effectively not considered (or are assumed equivalent).

4.5.3 Time horizon

The primary results of this cost-effectiveness analysis reflect the six month period of the trial. Any benefit (or harm) of the intervention was measured over a six month period. This approach implicitly assumes that any costs and effects in the

intervention and control groups occurring after the six month follow-up period are zero, and in this instance this may represent a conservative analysis (i.e. in favour of the control group). A sensitivity analysis examined the assumption that patients on the programme maintain the same EQ-5D score at 12 months as they have at the six month follow-up. As the intervention has been delivered, it is assumed that there is no difference in costs between the intervention and control groups in this hypothetical follow-up period.

4.5.4 Sensitivity analyses

Though the form of stochastic analysis described above addresses a large amount of uncertainty in the analysis, it is still appropriate to perform sensitivity analysis to allow for variability and methodological uncertainty. Four separate sensitivity analyses were performed; the first increasing the cost of the intervention costs to £450 per patient on the EPP, the second to include only costs falling on the health and community care sector and the third to reflect the possibility of the difference in utility between the groups being maintained over a longer period (in this case 12 months). The final sensitivity analysis used an alternative method, multiple imputation, to impute value for data that were missing at follow-up.

4.6 Results

The CONSORT diagram of the trial is presented below in Figure 4.3.

4.6.1 Missing data

A total of 629 patients were recruited to the trial between April 2003 and March 2005. There were no missing data on resource use or EQ-5D at baseline. However, there were missing data for both resource use and utility data at follow-up. At six month follow-up, 514 patients (82.0%) provided full EQ-5D responses; seven patients who completed the six month follow-up (as shown in the diagram above) did not complete the EQ-5D questionnaire. Of these 514, 243 were in the treatment group with 271 in the control group; thus missing values were higher in the intervention group than in the control group (22.1% vs 14.0%). 519 (82.7%) patients provided full resource use data, though again there were more missing values in the intervention group than in the control group (21.2% vs 13.3%).

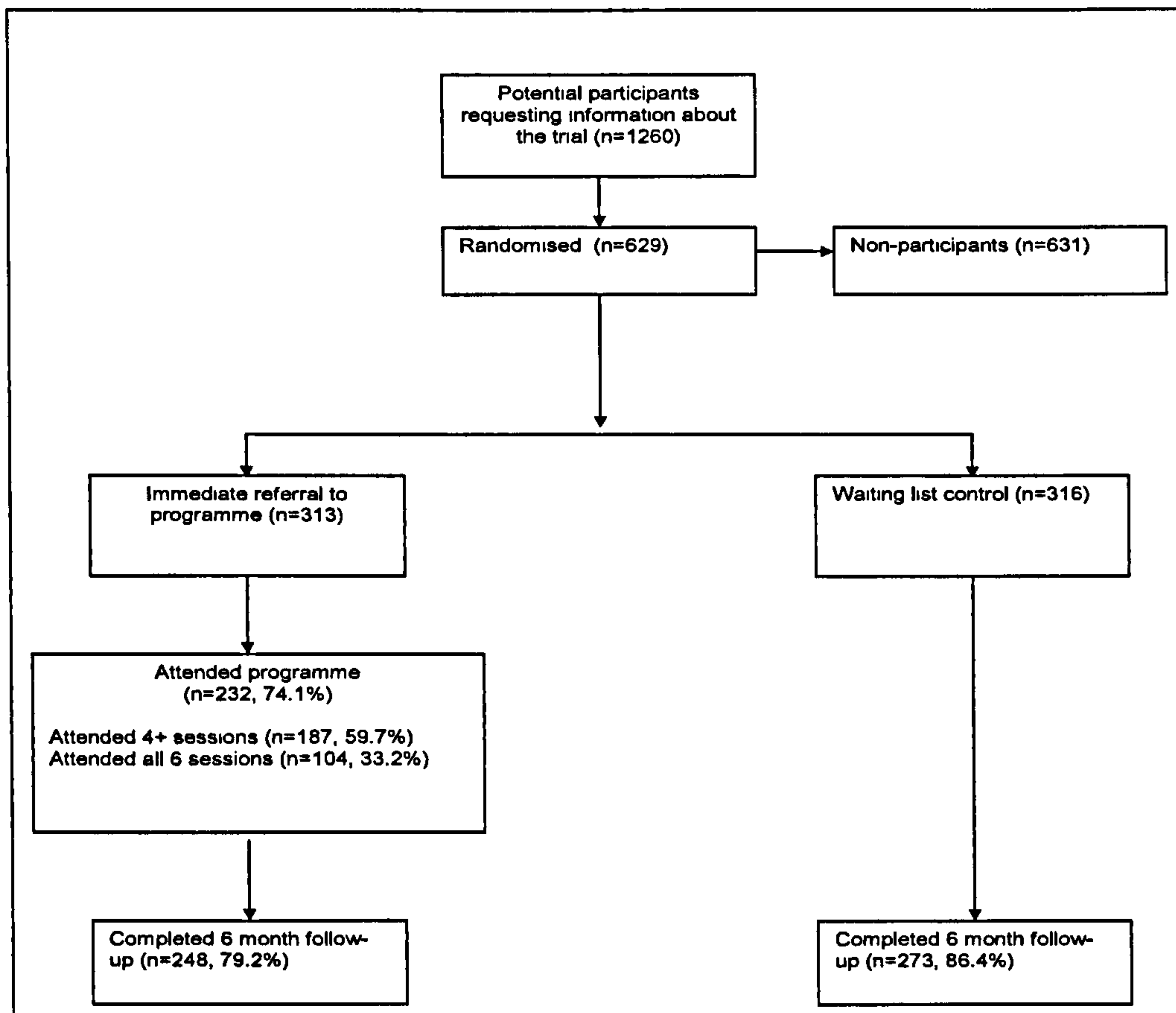


Figure 4.3 CONSORT diagram of the EPP national evaluation (reproduced from Kennedy et al⁶⁶)

4.6.2 Resource use

4.6.2.1 Unadjusted resource use

Unadjusted mean levels of resource use are presented below in Table 4.2. These estimates utilise resource use data estimated without the imputation method described above (i.e. are based on responders/completers of questionnaires).

For the majority of resource use variables above, the EPP programme resulted in a reduction in resource use. Though most of these differences are small, they are likely to offset some of the cost of the intervention.

Table 4.2: Mean resource use in the two groups over the 6-month period

	Intervention Group		Control Group		<i>Mean Difference (95% Confidence Interval)</i>
	<i>Mean</i>	<i>N</i>	<i>Mean</i>	<i>N</i>	
Inpatient length of stay	.80	246	1.59	272	-0.79 (-1.75 to 0.18)
Number of outpatient appointments	2.73	248	2.91	273	-0.18 (-1.17 to 0.81)
General Practitioner (at the surgery)	3.36	246	3.44	269	-0.08 (-0.65 to 0.49)
General Practitioner (at patient's home)	.09	247	.18	268	-0.09 (-0.18 to -0.01)
Practice Nurse (at GP surgery)	1.37	247	1.59	271	-0.22 (-0.77 to 0.32)
Occupational Therapist (at home)	.06	237	.22	264	-0.16 (-0.32 to -0.00)
District Nurse (at home)	.31	237	.23	264	0.08 (-0.34 to 0.52)
Home help	1.38	237	1.92	264	-0.54 (-2.32 to 1.25)
Meals on Wheels	.55	237	.02	264	0.53 (-0.56 to 1.61)
Physiotherapist (at home)	.11	237	.07	264	0.03 (-0.11 to 0.18)
NHS Direct	.32	237	.29	264	0.02 (-0.16 to 0.21)
Walk-In Centre	.08	237	.09	264	-.00 (-0.1 to 0.09)
Counsellor	.64	237	.60	263	0.04 (-0.46 to 0.54)

4.6.2.2 Adjusted resource use

In the EPP evaluation, patients were asked to estimate their use of resources over the previous six months (i.e before randomisation). For each item of resource use, this pre-baseline value was used as a covariate to adjust post-randomisation resource use.

The tables below show the results of the adjusted analyses. These analyses are based on the analysis of responders only (i.e. complete case analysis).

The second column of Table 4.3 replicates the analysis from Table 4.2 above. The third and fourth columns show analyses where the resource use estimates are adjusted for baseline differences. In the third column the resource use estimates are adjusted only for the treatment group and the relevant pre-baseline levels of resource use (for example GP visits are adjusted for pre-randomisation number of GP visits). In the final column, these are supplemented with adjustments for age, gender and condition. Adjusting for baseline covariates, though appropriate, has little impact on the results in most cases. The direction of the result changes in two cases, that of NHS Direct contacts and the cost of medication. In the former case, the unadjusted analysis showed the intervention was associated with a higher use, but when the estimates were adjusted for baseline covariates, the intervention was associated with a reduction in resource use. In the case of the cost of medication, the reverse is true. Initially, the unadjusted analysis demonstrated a reduction in drug costs associated with the intervention, but adjusting for covariates showed that there was in fact an increase in drug use associated with the intervention. This is likely to be due to the fact that treatment group is not the only factor influencing medication use. Thus the control group are likely to have had more “high medication use” individuals than the treatment group, and once this (and other factors) was allowed for in the analysis it becomes apparent that the effect of treatment group on medication is positive rather than negative. However, in absolute terms, both these changes are minimal and have little impact on the probability of the intervention being cost-effective.

Table 4.3: Mean resource use in the two groups over the 6-month period based on complete case analysis

	Unadjusted	Adjusted for baseline resource use	Adjusted for baseline resource use, age, gender, condition
	<i>Mean Difference (95% Confidence Interval)</i>	<i>Mean Difference (95% Confidence Interval)</i>	<i>Mean Difference (95% Confidence Interval)</i>
Inpatient length of stay (days)	-0.79 (-1.75 to 0.18)	-0.75 (-1.71 to 0.21)	-0.74 (-1.70 to 0.23)
Number of outpatient appointments	-0.18 (-1.17 to 0.81)	-0.06 (-1.01 to 0.89)	-0.04 (-1.00 to 0.91)
General Practitioner (at the surgery)	-0.08 (-0.65 to 0.49)	-0.03 (-0.54 to 0.48)	-0.02 (-0.53 to 0.50)
General Practitioner (at your home)	-0.09 (-0.18 to -0.01)	-0.10 (-0.18 to -0.02)	-0.11 (-0.19 to -0.02)
Practice Nurse (at GP surgery)	-0.22 (-0.77 to 0.32)	-0.22 (-0.77 to 0.32)	-0.23 (-0.79 to 0.32)
Occupational Therapist (at home)	-0.16 (-0.32 to -0.00)	-0.17 (-0.33 to -0.01)	-0.18 (-0.34 to -0.02)
District Nurse (at home)	0.08 (-0.34 to 0.52)	-0.02 (-0.44 to 0.38)	-0.03 (-0.44 to 0.38)
Home help	-0.54 (-2.32 to 1.25)	-0.32 (-1.28 to 0.64)	-0.37 (-1.33 to 0.59)
Meals on Wheels	0.53 (-0.56 to 1.61)	0.54 (-0.48 to 1.57)	0.53 (-0.51 to 1.56)
Physiotherapist (at home)	0.03 (-0.11 to 0.18)	0.03 (-0.11 to 0.18)	0.02 (-0.12 to 0.17)
NHS Direct	0.02 (-0.16 to 0.21)	-0.04 (-0.22 to 0.14)	-0.03 (-0.21 to 0.15)
Walk-in Centre	-0.00 (-0.1 to 0.09)	-0.00 (-0.1 to 0.08)	-0.00 (-0.1 to 0.09)
Counsellor	0.04 (-0.46 to 0.54)	0.06 (-0.34 to 0.46)	0.09 (-0.30 to 0.49)
Medication costs (£)	-23.57 (-174 to 127)	-0.06 (-108.4 to 108.3)	3.35 (-104.98 to 111.68)

4.6.3 Unit costs

Unit cost estimates and their sources used in the analysis can be found below in Table 4.4.

Table 4.4: Unit costs of resources used

	Unit Cost (£ 2003/4)	Source
GP cost per surgery visit	21.00	Curtis and Netten (2004)
GP cost per home visit	65.00	Curtis and Netten (2004)
Practice Nurse cost per visit	9.00	Curtis and Netten (2004)
District Nurse cost per visit	20.00	Curtis and Netten (2004)
Health Visitor cost per visit	31.00	Curtis and Netten (2004)
Home Help cost per visit	7.00	Curtis and Netten (2004)
Occupational Therapist cost per visit	48.00	Curtis and Netten (2004)
Physiotherapist cost per visit	48.00	Curtis and Netten (2004)
Counsellor cost per visit	33.00	Curtis and Netten (2004)
NHS Direct cost per visit	18.00	National Audit Office
Walk in centre cost per visit	25.00	Hansard
Meals on Wheels cost per meal	4.00	PSSRU Discussion Paper 1427
Cost per Inpatient attendance	Various (£203-£486)	Netten and Curtis (2002)
Cost per outpatient visit	Various (£63-£364)	Netten and Curtis (2002)
Cost of Intervention	£250 for the primary analysis	Department of Health ¹⁷²

4.6.4 Health States

Table 4.5. % of patients in each EQ-5D dimension by group at baseline and six month follow-up.

	EPP Group						Control Group					
	<i>% of patients in health state at baseline</i>			<i>% of patients in health state at 6 month follow-up</i>			<i>% of Patients in health state at baseline</i>			<i>% of patients in health state at 6 month follow-up</i>		
	1	2	3	1	2	3	1	2	3	1	2	3
Mobility	28.4	70.6	1.0	30.0	69.3	0.6	32.3	67.1	0.6	31.0	68.4	0.6
Self Care	59.7	38.7	1.6	62.9	36.1	1.0	61.7	36.7	1.6	56.0	42.4	1.6
Usual activities	22.0	67.1	10.9	22.0	72.2	5.8	22.8	65.5	11.7	23.1	68.4	8.5
Pain/discomfort	12.5	64.2	23.3	12.1	65.5	23.4	13.0	63.3	23.7	12.3	62.1	25.6
Anxiety/depress	37.1	52.1	10.9	44.7	50.8	4.5	32.9	58.5	8.5	37.3	53.8	8.9

For each dimension of the EQ-5D, there are 3 dimensions. 1 is the “best”, and represents “no problems” on that dimension, while 3 is the “worst” and represents extreme problems. For example on the mobility dimension, 1 represents “no problems walking about”, whilst 3 represents “confined to bed”.

While there were slight differences in the proportions in each category by group at baseline, none of these differences approached statistical significance, and do not appear to favour the intervention.

There appears to be little impact in either group on the mobility or pain dimensions. However, the other dimensions exhibit some noteworthy changes.

Both groups show an increased proportion in the least severe anxiety/depression, with the intervention group performing slightly better (in terms of the percentage in the least severe category). However, in the intervention group this is reflected by a reduction in the percentage in the most severely affected group, while in the control group, the reduction is mainly in the size of the moderately affected group. This demonstrates the importance of including an appropriate control group, simply comparing the “before and after” scores on these dimensions would overestimate

the treatment effect on this dimension (as individuals improved anyway without the intervention).

The self care dimension shows that the intervention group improved on average over the period, whereas the capabilities of the control group seemed to decline over the period. This was the only change that would be deemed “statistically significant”. Testing for difference in proportions shows that the proportion in category 2 (that is “some problems with self care”) in the control group is significantly higher than in the treatment group (two groups are 0.37 vs 0.28, with a difference of 0.090 and a 95% Confidence Interval around the difference of 0.87 to 0.93).

The usual care dimension showed a movement from the most severely affected to moderately affected, and this change was most marked in the intervention group.

Table 4.6: Mean EQ-5D score at baseline and follow-up by group

	Baseline	6 months
EPP group	0.470	0.605
Control group	0.474	0.538
Difference	-0.004	0.068
(95% CI)	(-0.056 to 0.048)	(0.015 to 0.120)

The changes shown in Table 4.5 are mirrored in the utility scores shown in Table 4.6. At baseline, there is little difference between the groups (though the control group have a slightly higher utility score). However, at six month follow-up, the EPP show a statistically significant increase in EQ-5D score at the conventional 5% level.

Another interesting result of this analysis is that *both* groups improve over the period of the study, which may appear surprising in a chronically ill population, again emphasising the importance of including a control group.

4.6.5 Health state and self-efficacy

Self-efficacy improved in both the treatment and control groups. However, the treatment group performed better on this measure than did the control group.

Individuals improved by 1.09 in the intervention group compared with 0.57 in the control (95% CI around the difference in means of 0.11 to 0.94). This improvement in both groups reflects the improvements in utility reported for both groups above. The relationship between the two outcome measures is displayed in Figure 4.4 below.

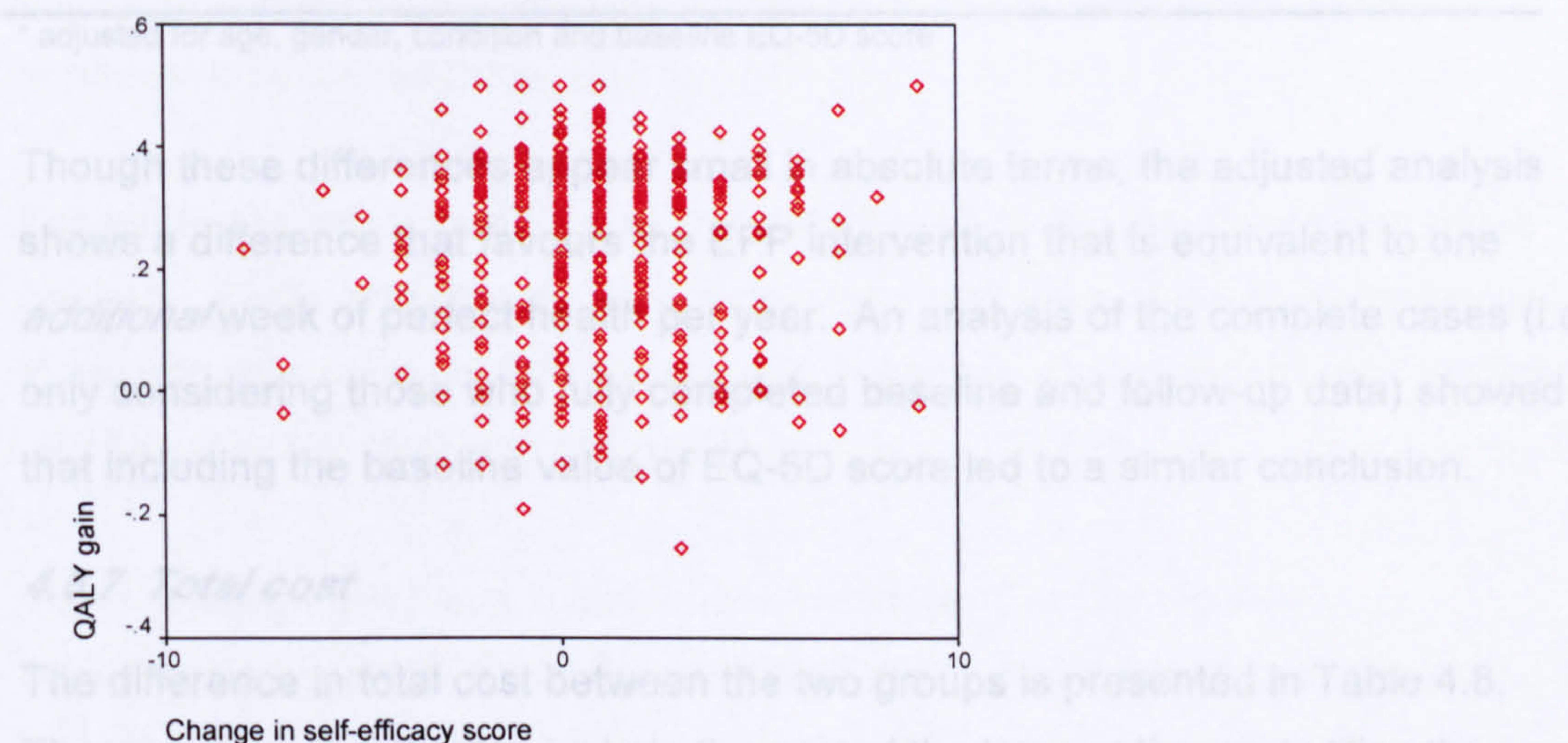


Figure 4.4 QALYs vs self-efficacy

There appears to be a positive relationship between the two outcome measures. Positive values of self-efficacy change are associated with positive values of QALYs (that is, there are a large amount of data points in the upper right hand section of the graph). However, it is worth noting that this is far from universal. There are a considerable number of data points that show improvements in self-efficacy associated with QALY losses and vice versa. The strength of self-efficacy as a surrogate outcome measure for health status is questionable based on these data. This issue is discussed in more detail in Chapter 6.

4.6.6 Quality-adjusted life-years (QALYs) based on imputed data

The EQ-5D score shows a difference between the groups at follow-up. Based on these estimates, the difference in total QALYs between the two groups can be estimated. The mean number of QALYs over the period is presented in Table 4.7. These estimates are based on data that include values imputed for missing values.

Table 4.7 Unadjusted and adjusted mean change in QALYs per patient over 12 month period

	Mean QALY	Difference (95% CI)	Difference allowing for baseline characteristics* (95% CI)
Intervention group	0.276	0.0184	0.020
Control group	0.258	(-0.004 to 0.041)	(0.007 to 0.034)

* adjusted for age, gender, condition and baseline EQ-5D score

Though these differences appear small in absolute terms, the adjusted analysis shows a difference that favours the EPP intervention that is equivalent to one *additional* week of perfect health per year. An analysis of the complete cases (i.e. only considering those who fully completed baseline and follow-up data) showed that including the baseline value of EQ-5D score led to a similar conclusion.

4.6.7 Total cost

The difference in total cost between the two groups is presented in Table 4.8. These summary estimates include the cost of the intervention and utilise the resource use estimates (adjusted for treatment group and baseline resource use) in Table 4.3. The total costs per patient are shown in the table.

It can be seen that when considering the impact on health care costs only, the EPP reduces costs considerably. However, while this may be of interest to health care decision makers, the societal cost is the relevant figure and this shows the EPP to be cost neutral.

Table 4.8 Total costs per patient in the two groups over 6 month period based on imputed data

	EPP group	Control group	95% CI around difference in mean cost
Health care costs only	£1169	£1560	-£389 (£38 to -£741)
Total cost including patient costs with Intervention costed at £250 per patient*	£1912	£1939	-£27 (£368 to -£422)

* the cost of the intervention is based on estimates from Department of Health calculated by dividing total cost of programme by throughput

There is again considerable uncertainty around these estimates and the difference is not statistically significant at conventional levels. Patients allocated to the EPP group have a slightly lower cost of providing health care (reflecting the resource use data presented in Table 4.2), though patient out-of-pocket costs are higher for the EPP group. Patient out-of-pocket expenditure included payments for alternative therapies, domestic help and special dietary needs as well as any home improvements that were necessary.

Table 4.8 presents a summary of cost data based on the results of the imputed analysis. Table 4.9 presents a more detailed breakdown of the costs by category. (Note that figures do not tally exactly with Table 4.8, as imputation leads to slight differences in the total cost estimates).

Other resource use items accounted for less than 1% of total costs.

As would be expected, these figures reflect the resource use data in Tables 4.2 and 4.3. However, it is noticeable that while the EPP programme reduces resource use (and costs) in most categories, patients' out-of-pocket expenditure rises in the EPP group. While the 95% confidence interval around the difference in mean costs includes zero, there is a suggestion that there may be some cost shifting from the NHS to individuals through improving patients' ability to self manage.

The major part of this difference between groups is explained by "housing alterations" (which accounts for £65 of the £115 mean difference), though the cost of "special dietary needs" and "complementary therapy" were also higher in the intervention group. The lack of a plausible mechanism for EPP to impact on housing alterations means that the impact of EPP on patient costs should not be overstated. Nevertheless, even allowing for this there is some evidence of increased out-of-pocket expenditure in the EPP group.

Table 4.9. Major cost items by group

<i>Resource type</i>	Intervention group mean cost (£ 2003/4)	Proportion of total costs	Control group mean cost (£ 2003/4)	Proportion of total costs
Inpatient stay(s)	247	12.9%	561	28.9%
Outpatient appointments	276	14.4%	294	15.2%
General Practitioner visits (surgery)	70	3.7%	72	3.7%
Medication costs	426	22.3%	450	23.2%
Patient out-of-pocket expenditure	493	25.8%	378	19.5%
Intervention costs	250	13.1%	0	0%
Counsellor visits	21	1.1%	20	1.0%
Day case	87	4.6%	127	6.5%
Subtotal*	1870	97.9%	1902	97%

* some categories omitted as individually below 1%

4.6.8 Incremental Cost-Effectiveness Ratio (ICER)

The EPP group is associated with an improved QALY profile and a slightly lower cost. Specifically, the EPP group have a 0.020 QALY gain compared with the control group, and a reduced cost of around £27 per patient. EPP would therefore be considered dominant and calculation of the ICER is inappropriate.

However, there is a large degree of uncertainty around these results. Therefore, to deal adequately with uncertainty the NMB approach was used and CEACs were generated. The results of these analyses are presented below.

4.7 Net Monetary Benefits and CEAC

The value of a QALY is open to some debate. However, the National institute for Health and Clinical Excellence (NICE) have suggested that interventions delivering a cost per QALY of under £20,000 are likely to be an acceptable use of NHS resources.¹² Figure 4.5 shows the CEAC curve. The cost-effectiveness threshold

(λ) is varied between zero (where gains in QALYs are not valued at all, and amounts to a test for differences in costs) and £50,000. In the main analysis with imputed data and a cost of the intervention of £250 and patient costs included, when the value of a QALY is £20,000 the EPP has a probability of 94% of being cost-effective. Indeed, for all plausible values of willingness-to-pay for a QALY the EPP group is more likely to be cost-effective than the control group. Figure 4.3 shows CEAC of the EPP for a variety of assumptions. Figure 4.4 presents a scatter plot of the incremental costs and effects of the EPP.

4.8 Sensitivity analysis

4.8.1 Higher intervention costs

The results of this analysis have been transformed into a NMB framework to enable the results to be presented in terms of the probability of the intervention being cost-effective, and allowing the uncertainty around the decision to be captured. These results show that under the assumptions of this analysis, with the intervention costing £450 rather than £250, the EPP group would be more expensive than control, with an ICER of about £8700 (using QALY gain adjusted for baseline covariates). This is represented in the CEAC by the lowest of the three lines. However, even with this assumption, the EPP has an 84% probability (at a willingness-to-pay of £20,000 per QALY) of being cost-effective based on these data.

4.8.2 Excluding patient costs

Patient costs were higher in the EPP group. From a societal perspective, it is important to include these costs, especially in the sphere of self care support where there could be a shift from care provision funded by government agencies to patients' out-of-pocket expenditure. A sensitivity analysis shows the impact of excluding these costs, and confirms that the EPP would be more likely to be cost-effective if patient costs were excluded, with a 97% probability of being cost-effective at a willingness-to-pay of £20,000 for a QALY. The calculation of an ICER is inappropriate as the EPP intervention is dominant under this scenario.

4.8.3 Maintenance of improvements

As discussed above, the main analysis is based on only 6 months data. It is likely that benefits (or harms) achieved within this period would be maintained for a longer period. The duration of this “maintenance” period is an empirical question, but this sensitivity analysis considered the impact of maintaining the EQ-5D scores from 6 months to 12 months. This sensitivity analysis additionally assumes that cost and QALY data were evenly distributed across the six month period. It is plausible that changes in costs or EQ-5D could have been incurred either very early or late in this period. However, there is no evidence to support these possibilities, and the assumption of equally distributed costs and EQ-5D was considered reasonable.

Using these assumptions, the EPP group has a 97% probability of being cost-effective at a willingness-to-pay of £20,000 for a QALY. Again, the ICER calculation is inappropriate due to dominance.

The assumption that costs would be the same between the two groups over the extended period is likely to be a conservative analysis as the post-intervention costs in the EPP group were lower than those in the control group. Therefore, extending the reduction in costs to 12 months would increase the probability that the EPP is cost-effective above 97%.

4.8.4 Multiple imputation

Missing data were initially imputed using best subset regression.¹⁷⁹ However, multiple imputation is an alternative methodology where multiple datasets are generated with potentially different values for each missing item imputed and gives a fuller reflection of the uncertainty surrounding which value to impute.¹⁷⁶ Multiple imputation was performed using the propensity score method in the software package Solas. Values were imputed for each of the dimensions of the EQ-5D (rather than total score), and for each missing item of resource use (rather than total cost). Five datasets were generated in this analysis. It is not clear how multiple imputation techniques can be used within the non-parametric bootstrap framework to generate CEACs; therefore the following estimates were generated

parametrically. The probability that the EPP intervention was cost-effective at a threshold value for a QALY of £20,000 in this scenario was 94%, despite the fact that the treatment group were slightly more costly (difference in imputed means of five datasets of £45) in this analysis. The ICER generated of £2,300 per QALY gained would be considered acceptable.¹²

4.8.5 Adjustment for condition, age and gender

The analysis above presents data where the costs were adjusted for pre-baseline differences in resource use. However, total cost data can be adjusted for many baseline covariates (assuming that patient level data are available). To illustrate, total cost data were adjusted for age, gender and condition as well as pre-baseline costs. In this case, the treatment group were approximately £37 less expensive than the control group (with a 95% CI around the difference in mean total costs of -£361 to £437). This is very similar to the unadjusted analysis with EPP considered dominant and calculation of the ICER inappropriate.

Of the covariates included, gender was the most influential, to the extent that if it were included the 90% CI would not include zero. Males cost on average £393 more than females (95% CI -£47 to £843).

The results of the sensitivity analyses above are presented in Table 4.10

Table 4.10 Sensitivity analyses

<i>Sensitivity analysis</i>	<i>Cost Difference (+ indicates EPP more costly)</i>	<i>QALY Difference (+ indicates EPP more effective)</i>	<i>ICER (£)</i>
Higher intervention costs	+175	+0.02	8,700
Exclusion of patient costs	-140	+0.02	Dominant
Extending time horizon	-27	+0.04	Dominant
Multiple imputation	+45	+0.02	2,300
Adjustment for covariates	-37	+0.02	Dominant

4.8.6 Conventional (classical) statistical analysis

All of the analyses above employ Bayesian inference techniques. It was argued in Chapter 2 that this is the appropriate methodology for economic evaluation to inform decision making and results in an estimate that the EPP is 94% likely to be cost-effective given a threshold value of £20,000 for a QALY. A frequentist analysis could test for significant differences, usually at the 5% level of significance, in cost and effect. At these levels, there is no significant difference in cost, but a significant difference in effect when adjustments are made for baseline differences. Clearly, which variables are included in the analysis also involves a degree of subjectivity, but it is considered appropriate to adjust for differences in age gender, condition and baseline EQ-5D score.⁴⁴ In this instance, as there is no significant difference in costs but a significant improvement in patient outcomes, the frequentist interpretation would concur with the Bayesian analysis that the EPP is a cost-effective use of resources.

4.9 Discussion

The analysis above shows that, for most reasonable threshold values of a QALY, the EPP would be considered likely to be cost-effective when compared with treatment as usual. This conclusion that the EPP is cost-effective is maintained whether the analysis is conducted in the Bayesian or frequentist paradigm, and also over a range of plausible assumptions.

The higher probability of EPP being cost-effective compared to control is due to there being little difference in costs and an improvement in QALY scores in the EPP group. Costs in both groups were similar as the increased patient costs together with the cost of delivering the intervention was offset by reductions in resource use.

However, there is considerable uncertainty around the estimates of costs and effects in each group. This uncertainty results in a probability of between 84% and 97% (depending on the assumptions made) that the intervention is cost-effective (and therefore a probability of between 3-16% that not providing EPP is cost-effective, as only one comparator was considered).

The time horizon of the study was limited to only six months (one year in a sensitivity analysis). Clearly, in this patient population, there is the potential for longer term benefits. Ideally, a longer term trial would have been conducted but pragmatically this was not possible. It is likely that a longer trial would have increased the probability of EPP being cost-effective (as shown in the hypothetical one year model), subject to the caveats expressed in section 4.5.3.

The exclusion of patient costs increased the probability of the EPP being cost-effective. In this instance, the decision would not be changed (the EPP is likely to be cost-effective whether these costs are included or excluded), though this may not always be the case and particularly in evaluations of interventions designed to support self care, these costs should be included.

The analysis above shows the potential importance of adjusting for baseline covariates. The direction of the results on some cost items (such as medication) changes from being in favour of the intervention to being in favour of the control, while the magnitude of the QALY difference is increased when baseline variables are included.

Previous evaluations of self care support interventions were described in Chapter 3. The most common limitations of these trials were that they had poor costing methodology (often excluding patient costs), inappropriate comparison groups, inadequate handling of uncertainty, inappropriate methods of analysis (often not intention-to-treat) and usually a short follow-up period. Ideally, the problems encountered in previous studies should be used to inform the design of new studies. In this evaluation, some of these issues were addressed. In particular, the costing methodology was much more thorough than in most previous studies and included the potentially important inclusion of patient out-of-pocket expenses. The handling of uncertainty and the methods of analysis were also more appropriate than in some of the previous studies where missing data were often ignored and/or uncertainty around estimates was not presented. However, the control group in the EPP evaluation was a waiting list control. It would be preferable to have another therapy (such as another group therapy) where the additional effectiveness of the intervention could be established in isolation from the effectiveness of being in a

group of similar chronically ill patients. In this case, however, the intervention was already being rolled out across England. The potential to randomise patients to another intervention when the EPP was available would increase the likelihood of contamination.

This evaluation of the EPP also had a short follow-up period (six months data with randomisation intact) due to the requirement by the funder for expeditious results. Additional analyses were performed to address the concerns with the short follow-up and to give better estimates of the long-term effects of the EPP. Extending the time horizon of the study may favour of the EPP intervention, though this depends on the timing of the costs and benefits; this is discussed in more detail in section 5.4.1.2.

The improvements in patient outcomes as measured by EQ-5D showed that patient's ability to self care increased in the treatment group as compared with control. This is intuitively appealing as the EPP has improved individuals abilities to self care, and is also consistent with the theoretical model whereby improvements in self-efficacy (demonstrated in the clinical trial),⁶⁶ lead to an improvement in health status.

The generalisability of these results is open to some debate. Around 70% of the sample were female (consistent with other studies of this intervention) and 95% were of "white" ethnicity, clearly raising the question of how cost-effective the intervention would be in ethnic minority (or more male) populations. Indeed, the other UK based trial was conducted in a population of Bangladeshi people, and the impact on QALYs was much less marked, and the cost-effectiveness much less certain. These issues are investigated in Chapter 5.

4.10 Conclusion

The results of this single trial based economic evaluation largely favour the introduction of the EPP, with a slight reduction in cost and a small but important improvement in QALYs. These results reflect those of the clinical trial⁶⁶ that demonstrated a large (and statistically significant) impact on self-efficacy, the

primary outcome of the clinical study. There was also a significant improvement in energy, which in the clinical study was used as a measure of health status.

There are other data that could be included in the exploration of the cost-effectiveness of the EPP. One other economic evaluation of the EPP has been carried out in the UK,⁸³ and the synthesis of these two datasets is explored in the next chapter. In addition, there are previously published evaluations of CDSMP, performed in the US and elsewhere. These show that self-efficacy is enhanced by this programme, though there are limitations in the analysis of these previous studies. Chapter 6 extends the analysis to include the impact of these studies.

While this chapter presents the evidence from one economic evaluation conducted alongside a randomised controlled trial, later chapters will synthesise the evidence from this trial with other “relevant” evidence. The incorporation of this additional evidence extends the amount of studies included, but there is a trade-off as the relevance to the UK decision context is reduced. In addition, the incorporation of additional evidence requires stringent assumptions about the relationship between treatment effect and outcomes to be met. This is discussed more fully in Chapter 6.

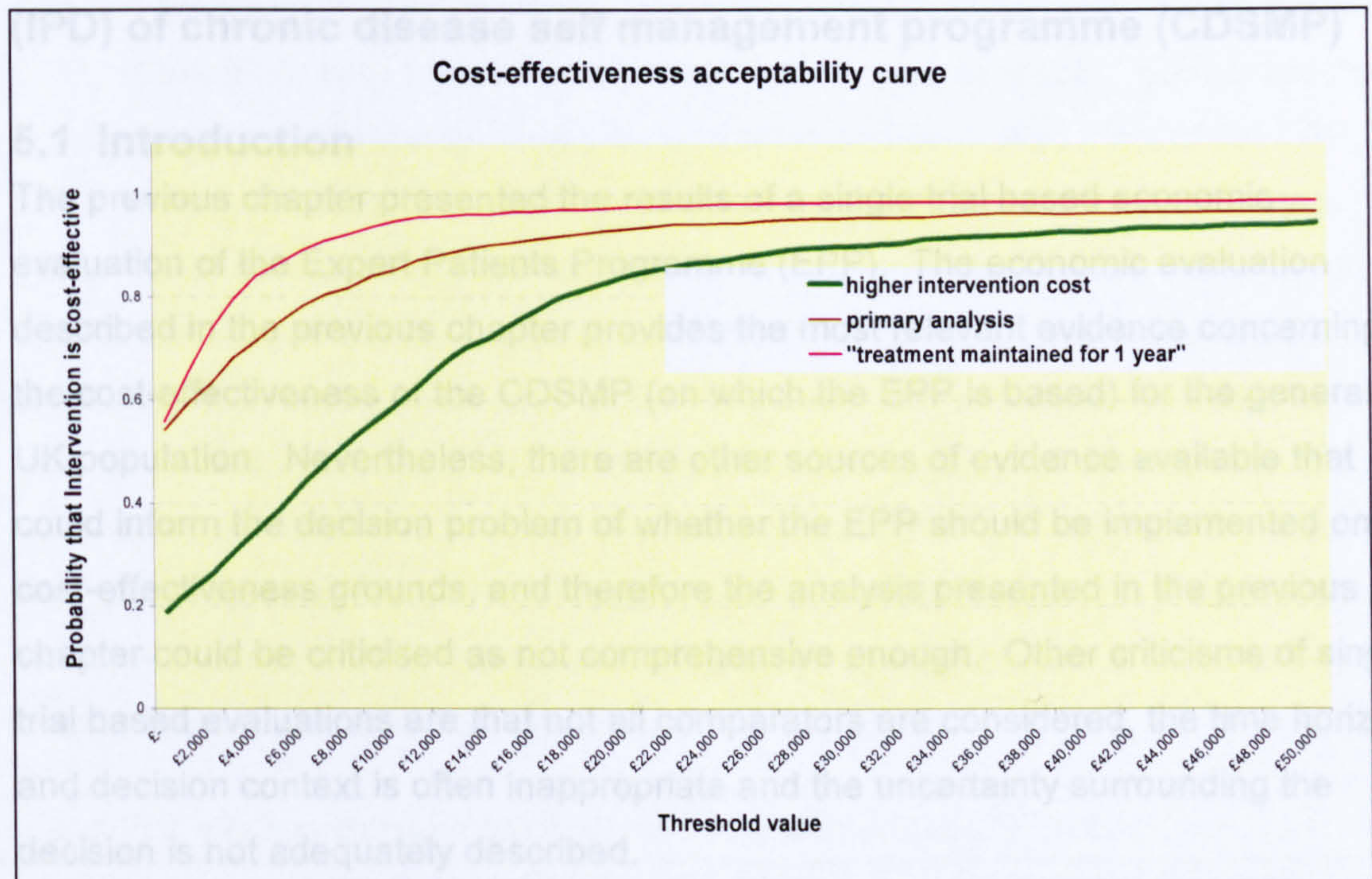


Figure 4.5 EPP Cost-effectiveness acceptability curve

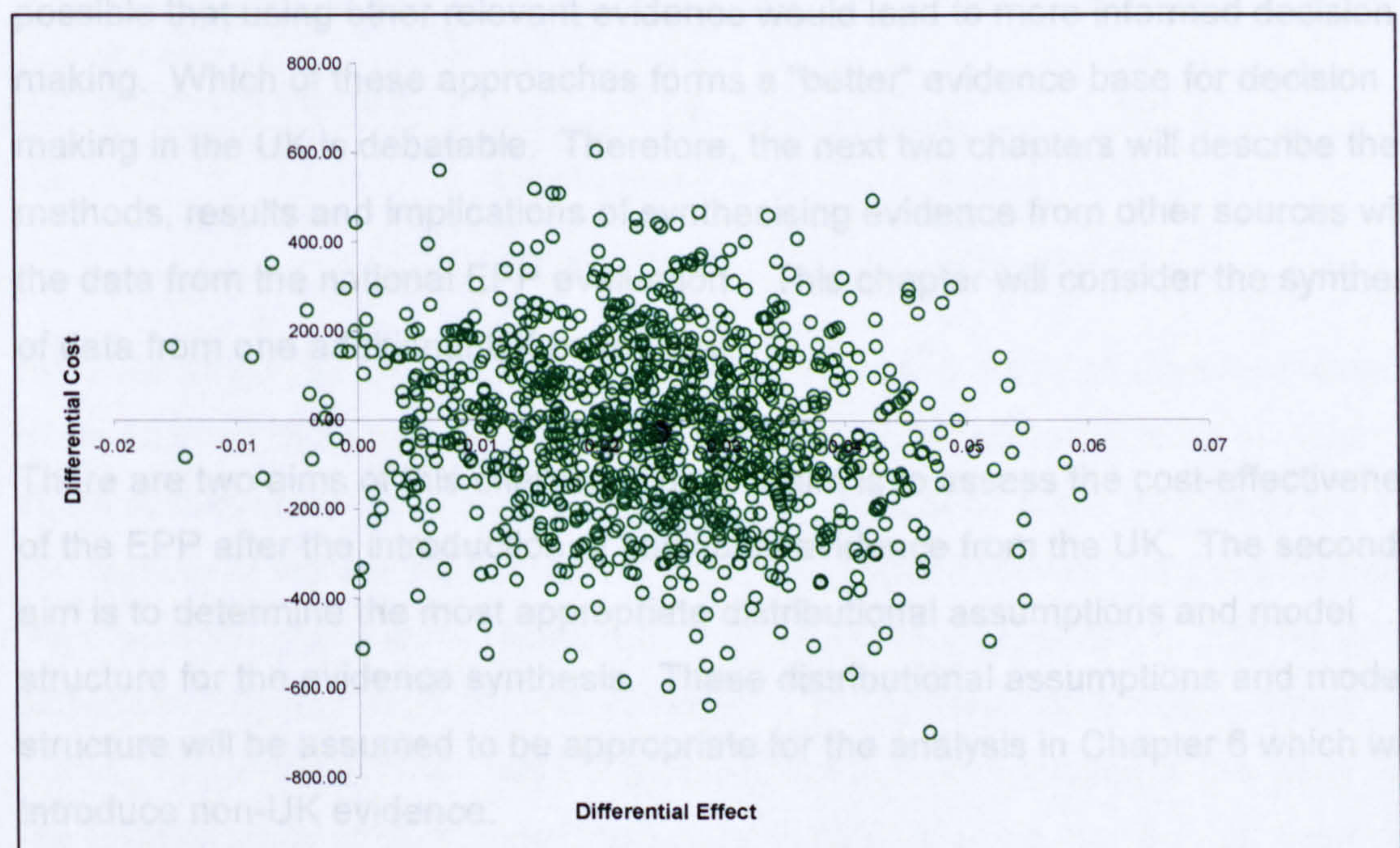


Figure 4.6 Scatter plot of incremental costs and effects of EPP

Chapter 5. Evidence synthesis of trials with Individual Patient Data (IPD) of chronic disease self management programme (CDSMP)

5.1 Introduction

The previous chapter presented the results of a single trial based economic evaluation of the Expert Patients Programme (EPP). The economic evaluation described in the previous chapter provides the most relevant evidence concerning the cost-effectiveness of the CDSMP (on which the EPP is based) for the general UK population. Nevertheless, there are other sources of evidence available that could inform the decision problem of whether the EPP should be implemented on cost-effectiveness grounds, and therefore the analysis presented in the previous chapter could be criticised as not comprehensive enough. Other criticisms of single trial based evaluations are that not all comparators are considered, the time horizon and decision context is often inappropriate and the uncertainty surrounding the decision is not adequately described.

For the evaluation of the EPP, it was argued in Chapter 4 that the use of single trial based cost-effectiveness analysis (CEA) *may* be appropriate. However, it is also possible that using other relevant evidence would lead to more informed decision making. Which of these approaches forms a “better” evidence base for decision making in the UK is debatable. Therefore, the next two chapters will describe the methods, results and implications of synthesising evidence from other sources with the data from the national EPP evaluation. This chapter will consider the synthesis of data from one additional UK based trial.

There are two aims of this chapter. The first aim is to assess the cost-effectiveness of the EPP after the introduction of additional evidence from the UK. The second aim is to determine the most appropriate distributional assumptions and model structure for the evidence synthesis. These distributional assumptions and model structure will be assumed to be appropriate for the analysis in Chapter 6 which will introduce non-UK evidence.

5.2 Trial characteristics

The two trials included in the analysis were the only UK based evaluations of the EPP. In addition, both trials had Individual Patient Data available. Both of the trials included in this evidence synthesis were randomised controlled trials of the EPP, with a waiting list control. That is, patients who were randomised to the control group would get the intervention at the end of a pre-specified period. This period differed between the trials requiring adjustments to be made; these are detailed in section 5.4.1.2, below. The aims of the trials were both to assess the effectiveness and cost-effectiveness of the EPP, though the patient groups that were targeted were different. Thus, the national evaluation of the EPP^{66, 162} targeted a sample that was representative of the UK population, while the Griffiths trial⁸³ clearly aimed to consider those from a specific ethnic minority. Characteristics of patients involved in the trials are presented in Table 5.1.

Table 5.1 Comparison of baseline characteristics of two IPD trials

	<i>National evaluation of EPP^{66, 162} (n=629)</i>	<i>Evaluation of EPP in UK based Bangladeshi population⁸³ (n=576)</i>
Mean Age (standard deviation)	55.4 (13.6)	48.9 (9.9)
% female	70	57
% from ethnic minorities	5	100
% in employment	19	9
Marital status (% married)	60	86
"Main" chronic condition (%)	<i>Other 49 Musculoskeletal 34 Diabetes 11 Cardiovascular 5</i>	<i>Diabetes 69 Asthma 16 Arthritis 10 Cardiovascular 6</i>

Not surprisingly, as the aim of the Griffiths study was to establish the effectiveness of the CDSMP in a minority South Asian population, the characteristics of individuals in the studies were very different. The national evaluation of the EPP recruited a low number of individuals from ethnic minorities, though this was only slightly below the proportion of ethnic minorities with chronic conditions reported in the Health Survey for England at 5% compared with 6%.¹⁸⁰ Age was, on average, higher in the national evaluation, as was the proportion of females recruited to the study. In the Griffiths study sample, individuals were more likely to be married, but were less likely to be employed, though both studies showed that employment rates were low in those with chronic conditions.

5.2.1 Trial results

The national evaluation⁶⁶ concluded that the introduction of EPP was likely to be cost-effective. In summary, this evaluation showed an improvement in QALYs associated with the treatment group, with no additional costs. In the Griffiths study in a Bangladeshi population, there was little impact on either resource use or QALYs, though additional costs of the intervention meant that the intervention is unlikely to be cost-effective in this population. Both these analyses use baseline EQ-5D as a covariate as recommended in the literature.⁴⁴ It is interesting to note that this allowance for baseline EQ-5D changes the direction of the result in the Griffiths study. Without the baseline adjustment, the group receiving EPP performed slightly better than the control (generating an additional 0.006 QALY per person on average); however, the appropriate analysis including this covariate showed that the intervention group actually performed slightly worse than the control group, as shown in Table 5.2 below.

Table 5.2 Incremental costs and QAL Ys over 12 months in the two trials

	<i>National Evaluation of EPP (95% CI)</i>	<i>Griffiths evaluation of EPP (95% CI)</i>
Incremental QALYs (intervention minus control, adjusted for baseline EQ-5D)	0.020 (0.007 to 0.034)	-0.002 (-0.014 to 0.012)
Incremental cost (intervention minus control)	-£27 (-£422 to £368)	£146 (£65 to £223)

So the meta-analysis will synthesise evidence from two trials, one showing the intervention to be *dominant*,^{66, 162} the other showing that it is *dominated*.⁸³

5.3 IPD meta-analysis

IPD were available for both trials and IPD meta-analysis of cost-effectiveness data can therefore be carried out. Performing this analysis incorporates more relevant information, generates a more accurate assessment of the uncertainty around the decision and broadens the decision context by incorporating data from less advantaged groups. In addition, the use of IPD meta-analysis has several methodological advantages compared to the use of summary statistics. In

particular, it enables the inclusion of patient level covariates to “adjust” for certain characteristics (such as age, gender, baseline utility levels) in CEA if this is appropriate. This is not possible with aggregate data where typically a mean and standard error are the only data presented.

In addition the use of IPD meta analysis can avoid biases that can be apparent in meta-regression using summary statistics.¹⁸¹ Potentially dangerous biases that have been identified in meta-regression include publication bias and ecological fallacy.¹⁸² The former is known to occur due to studies with more significant, “interesting” results or that are of higher quality and are more likely to be published,¹⁸³ leading to potentially biased conclusions. Though this is an issue with meta-regression, it is worth noting that this is also a potential danger with IPD meta analysis. IPD meta-analysis involves the identification of trials and successfully accessing the IPD of these trials. If trials are only identified because their results have been published then there remains a possibility that there were unpublished trials with IPD available. Nevertheless, with the proliferation of trial registers this is, perhaps, becoming less of an issue.

Ecological fallacy occurs where the association that is observed between aggregate level variables is not a reflection of the association at the level of the individual, that is, correlations between variables at an aggregate level may not be comparable with correlations between variables at an individual level. An example of this is the impact of age on treatment effect. Plotting the relationship between mean age and treatment effect from several trials may show, for example, that the treatment effect appears to diminish with age. However, examination of individual data within each trial may show that there is in fact a positive relationship and that treatment effect improves with age. This may be due to the fact that the trials differ in some other respect, for example the trials with higher average age and lower estimates of treatment effect may have recruited healthier individuals and it is this (confounding) variable that the across trial relationship detects. This has led to the recommendation that within trial IPD should be used to estimate the relationship between “treatment benefit and patient characteristics, so that confounding because of differences across trials is avoided”.¹⁸⁴

Another advantage of IPD meta-analysis, and one that is particularly important for this analysis, is that cost-effectiveness analyses can be conducted that are consistent across studies.¹⁸¹ When synthesising data from different cost-effectiveness studies, there are many potential inconsistencies. For example, costs may be calculated using different unit cost data, different assumptions about future costs and benefits may be made, the perspective of the studies could be different (for example NHS versus society) and so on. In addition, there is some evidence that QALYs are not calculated consistently in cost-effectiveness analysis and that alternative covariates could be used in the estimation of incremental QALYs.⁴³

These issues are dealt with in more detail in section 5.4.

5.4 Methods

Bayesian analysis is particularly useful for evidence synthesis in that it enables all the uncertainty (including the uncertainty arising from the existence of missing data) and correlation from heterogeneous trials to be incorporated.⁸⁴ Nevertheless, where there are obvious sources of heterogeneity, such as different follow-up periods, these can be adjusted before synthesis (see section 5.4.1.2). At present there are two trials with IPD available, and the following section describes the methods required to meta-analyse these data. This analysis takes a Bayesian perspective where it is necessary to treat parameters as random variables and generate posterior distributions (or final estimates) for these parameters based on:

- a) The model or estimation procedure. How samples are generated depend on the formulae and distributions depicted in the model
- b) The choice of initial values.
- c) The “priors” and the assumptions around them

The models and estimation procedure are dependent on the choice of model that is developed in section 5.4.3 (that is, which covariates are included and the distributional assumptions around the costs and effect estimates). The generation of priors and initial values, model specification and number of required iterations as well as the procedures for ensuring the model has behaved appropriately are

described in section 5.4.3. Sensitivity analysis around the model assumptions are described in section 5.4.4.

The first sections, sections 5.4.1 and 5.4.2, describe adjustments required to ensure that the datasets are comparable and analyses are consistent.

5.4.1 Trial characteristics

Before meta-analysis can be properly conducted, it is necessary to examine individual trials to ensure their comparability to enable consistent analysis to be conducted. The characteristics of the trials and the patients within them have been considered in section 5.2.

5.4.1.1 Similarities between trials

Full details of both trials can be found elsewhere.^{66, 83} Briefly, both trials were randomised controlled trials (RCTs) with the patient as the unit of randomisation and also the unit of analysis. In addition, both trials had “treatment as usual” as the comparator and had relatively short follow-up periods and therefore discounting was not appropriate as all costs within both trials fell within a one-year period.

5.4.1.2 Ensuring trials are comparable and analyses are consistent

Despite some similarities, there were differences between the methodologies employed in the trials. Specifically, the Griffiths trial⁸³ used the perspective of primary and secondary care services within the NHS, while the EPP national evaluation employed a wider perspective including community care costs (such as district nurse visits, NHS direct contacts, physiotherapist and occupational therapist contacts), drug costs and patient out-of-pocket expenditure. In addition, the Griffiths trial had 4 months follow-up, while EPP national trial had 6 months follow-up.

To ensure comparability of the two studies, the costs not directly estimated in the Griffiths study were estimated (by group) from the EPP national trial and then simply added to the Griffiths data. Clearly this makes the assumption that these costs are exchangeable. However, in the absence of any evidence to the contrary or any

better estimates of the additional expenditure, this is considered more appropriate than excluding costs which may be relevant to the decision.

Also, to ensure the follow-up period was comparable and that analysis was consistent, the data from the former trial were extrapolated from four to six months by multiplying both QALYs gained and costs by 1.5. These “new” cost data were then appended to the EPP trial data ⁶⁶ using the statistical package Stata 8. This makes the additional assumption that the cost and QALY data were evenly distributed across the trial period. While there is no evidence to contradict this assumption, it is plausible that costs (or resource use) could have been incurred mainly in the first few weeks and that the extension of the trial period from 4 to 6 months would have little impact on overall costs. Similarly, it is plausible that the QALYs generated were distributed largely in the first part of the follow-up period. In the absence of evidence to support or refute any of these possibilities, the decision to assume an equal distribution of costs and QALYs across the follow-up period was considered reasonable.

5.4.2 Meta-analysis of IPD

There are several methods available for the meta-analysis of IPD. Simmonds et al reviewed published methodologies in the meta-analysis of IPD and identified three strategies:¹⁸¹

- i) “Mega-trial”. Data are analysed as if they were from a single trial
- ii) Stratified analysis. Trial identifier is included in the model using either a fixed or random effects model.
- iii) Two-stage approach. Summary statistics of each trial are produced and then combined as in a standard meta-analysis.

Each of the above methods was considered. The software package WinBUGS ⁹⁷ was employed for all meta-analyses.

5.4.2.1 “Mega-trial”

In this approach, data are treated as though they were from the same trial.

This approach assumes perfect exchangeability between trials (after adjustment for the above methodological differences).

Clearly, the mega-trial approach is a crude approach to synthesising IPD. The underlying assumption of this approach is that the data are perfectly transferable. For effectiveness data in the same geographical area using patients with similar characteristics, there may be some merit in this approach, though it is unlikely to be a correct assumption for cost data. This simple model can be extended by using patient level covariates as explanatory variables. However, due to the crudity of the approach and the necessary loss of useful data, this method was not considered further.

5.4.2.2 Fixed effect analysis

A preferable approach to the meta-analysis of IPD uses the trial identifier in the model. This allows differences between trials to be included and is more realistic than the mega-trial presented above. These “stratified” analyses can be performed either using a fixed effect or random effect. With the fixed effect model, each trial within the model has a different baseline. Essentially there is a relative treatment effect common to both trials but each trial has a specific constant (equivalently the estimated regression line for each trial is raised/lowered by a specific amount). Fixed effects models are extended in the analysis below by incorporating patient level covariates in the analysis. Random effects models were deemed inappropriate in this analysis as there were only two datasets of IPD. Attempting to establish a distribution around the differences between trials in the treatment effect based on two observations is futile.

5.4.2.3 Two-stage approach

The two-stage approach essentially creates aggregate estimates of costs and effects for each trial and then uses standard meta-analysis techniques on these aggregate estimates. In this methodology, there is less scope for using individual patient characteristics as covariates. Where IPD data are available, this form of analysis excludes relevant information.¹⁸⁵ In this instance, as IPD were available, this analysis was not conducted.

5.4.3 Modelling assumptions

Models were run using the statistical package R, and the simulation package WinBUGS. Full details of the code in both packages are presented in Appendix C.

5.4.3.1 Cost of intervention

The intervention cost in the EPP national trial was £250 and included a full costing including volunteer time, hiring of venues and equipment as well as the costs of running the programme. The Griffiths trial estimated a lower intervention cost of £123. The intervention cost used in this analysis is a simple weighted mean (weighted by sample size in each trial) of the two estimates.

5.4.3.2 Fixed effect models

Section 5.4.2 described alternative model structures using a mega-trial approach, the two stage approach, or the preferable fixed effect model where differences between the two trials are acknowledged. The analysis concentrated on fixed effects models which can be represented as:

$$C = \alpha + \beta_{C_T}Tx + \beta_{C_S}study + \beta_{C_X}X_i + \epsilon \quad \text{Eq 1}$$

$$Q = \alpha + \beta_{Q_T}Tx + \beta_{Q_S}study + \beta_{Q_X}X_i + \epsilon \quad \text{Eq 2}$$

where Tx is treatment group

β_{C_T} and β_{Q_T} is the effect of treatment on costs and QALYs respectively

β_{C_S} and β_{Q_S} is the effect of study on costs and QALYs respectively

β_{C_X} and β_{Q_X} is the effect of other variables on costs and QALYs respectively

and ϵ is the error term

In addition, it is plausible that the covariates identified above may have different impacts in the two studies. Thus study specific covariates were also considered. If, for example, baseline EQ-5D was included as a covariate in the cost regression, but that the relationship between costs and baseline EQ-5D might differ between studies, the models would then be represented as:

$$C = \alpha + \beta_{C_T}Tx + \beta_{C_S}study + \beta_{int}study*EQ-5D + \beta_{C_X}X_i + \epsilon \quad \text{Eq 3}$$

where β_{int} is the study specific impact of EQ-5D on costs

Clearly, a similar approach can be used for QALY estimation. The inclusion of specific study effects is different from the inclusion of covariates. In the latter case we have little or no *a priori* expectations of the impact these covariates have on costs and/or effects. However, in the case of study level effects we do have reason

to believe that costs (particularly) and effects may differ between geographical area with different accounting methods and in different patient populations. The other covariates that may be appropriate for inclusion are discussed below

5.4.3.3 Choice of covariates

The importance of allowing for baseline differences between treatment and control groups is becoming increasingly recognised.⁴⁴ However, it is necessary to determine which covariates should be included in models. Ideally, models with and without particular covariates could be run in WinBUGS and subsequently assessed for model fit. Unfortunately, running such models is time consuming, so at present it is necessary to simplify the model structure by identifying the appropriate covariates for inclusion using the regression techniques described below. The regressions were run in R. Subsequently, WinBUGS is used to identify the appropriate distributional assumptions of costs and benefits within the model (see section 5.5.3.3).

For the purposes of this analysis, we were interested in the effect of treatment on the outcomes of interest (costs and QALYs). Additional differences between treatment and control groups may have been influenced by a number of candidate covariates. These included age and gender of participants, their baseline levels of self-efficacy and their baseline health related quality of life as measured by EQ-5D. Unlike the study effects discussed in section 5.4.3.2, there was no *a priori* reason for any of these covariates to be given precedence over the others in that all have a plausible link with both costs and QALYS. All of these variables were therefore included in a regression model. Thus the initial models for costs and QALYs in the fixed effect framework were as follows:

$$C = \alpha + \beta_{C_T}Tx + \beta_{C_S}study + \beta_{C_A}age + \beta_{C_G}gender + \beta_{C_SE}SE + \beta_{C_E}EQ-5D + \varepsilon \quad \text{Eq 4}$$

$$Q = \alpha + \beta_{Q_T}Tx + \beta_{Q_S}study + \beta_{Q_A}age + \beta_{Q_G}gender + \beta_{Q_SE}SE + \beta_{Q_E}EQ-5D + \varepsilon \quad \text{Eq 5}$$

where Tx is treatment group, SE is self-efficacy

β_{C_T} and β_{Q_T} is the effect of treatment on costs and QALYs respectively

β_{C_S} and β_{Q_S} is the effect of study on costs and QALYs respectively

β_{C_A} and β_{Q_A} is the effect of age on costs and QALYs respectively

β_{C_G} and β_{Q_G} is the effect of gender on costs and QALYs respectively

β_{C_SE} and β_{Q_SE} is the effect of self-efficacy on costs and QALYs respectively

β_{C_E} and β_{Q_E} is the effect of EQ-5D on costs and QALYs respectively

and ε is the error term

Stepwise regression with backward selection was performed in R using these regressions separately for costs and QALYs. This involves starting the model with all potential covariates included. The variables are then tested individually for statistical significance and the variable that is the least significant is removed.

The model was then re-run until the only variables remaining were those with a statistically significant coefficient. The resulting model was judged as appropriate. The Akaike Information Criteria (AIC) was also monitored. The AIC improves as goodness of fit improves, but there is also a penalty for increasing the number of parameters in the model. Thus a model which explains more of the variation in the data may be worse (on AIC) if it uses more parameters to explain this variation. The lower the AIC value, the better is the model on this criterion.

5.4.3.4 Distributional assumptions

The use of log normal distributions in the software package WinBUGS is more computationally time consuming than using normal distributions. Therefore, where data are approximately normal this is the most efficient choice of distribution. However, cost data are rarely normally distributed. Examination of the plots below shows that a log normal distribution on costs may be a more appropriate assumption than normally distributed costs. Using log normal distributions requires that there are no zero values. In this instance, only one patient incurred a zero cost and this was altered to a cost of £1 to enable simulation to proceed (as the log of zero is undefined).

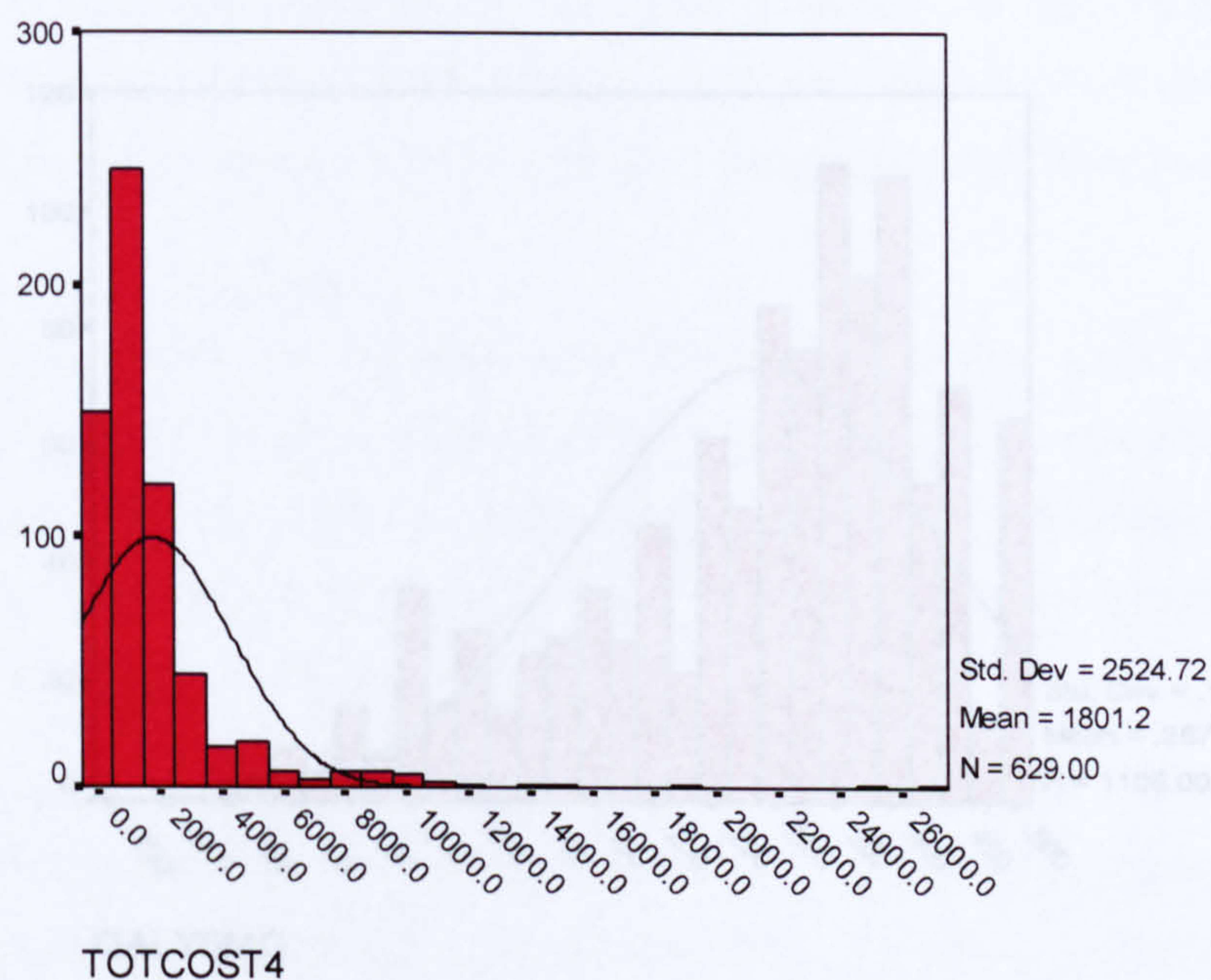


Figure 5.1. Histogram of untransformed total costs of EPP and Griffiths combined

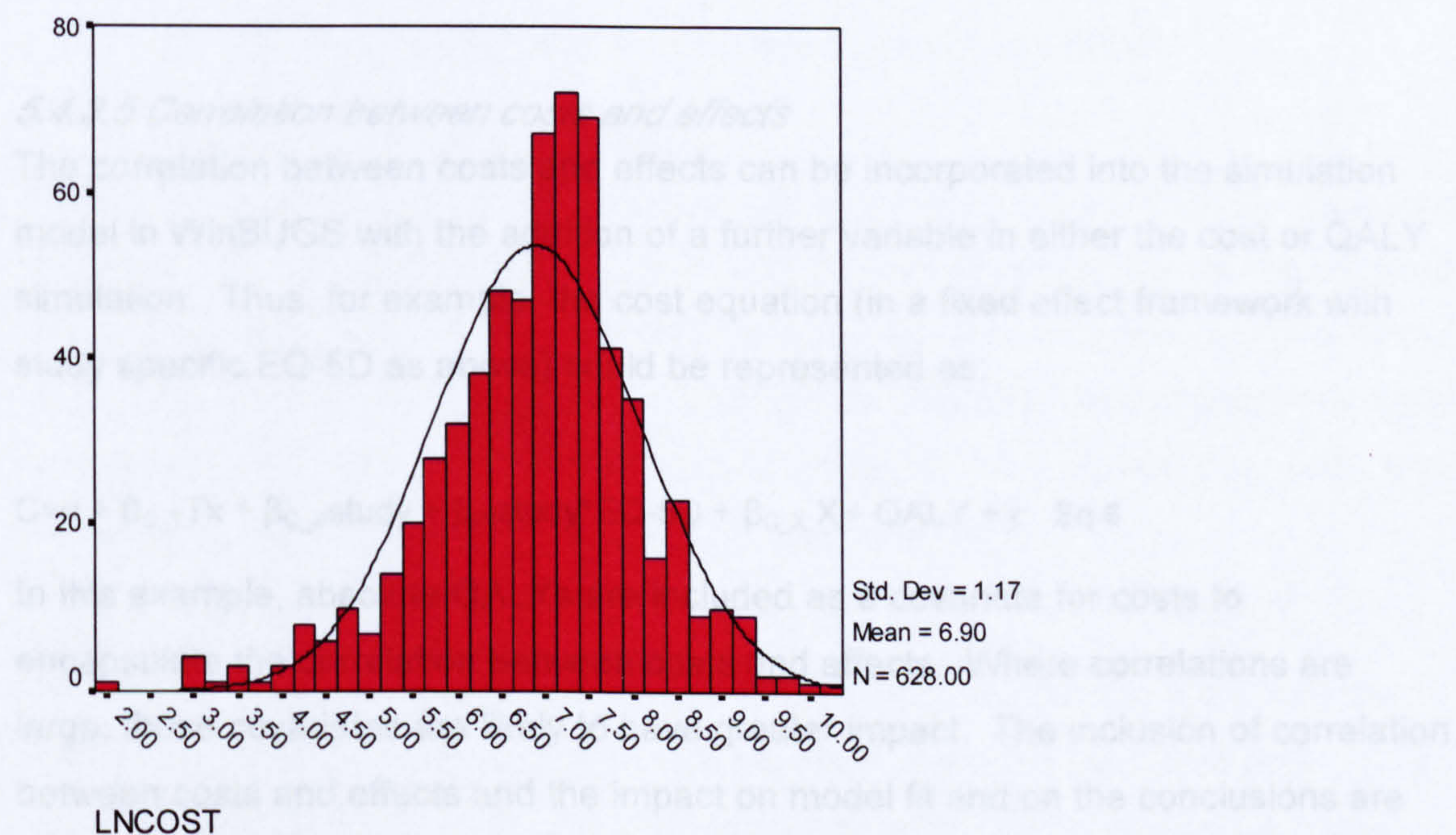


Figure 5.2. Histogram of (natural) log costs

The distribution of QALYs is presented in Figure 5.3 below. Though the data are skewed slightly to the right, a normal distribution was deemed appropriate in the analysis.

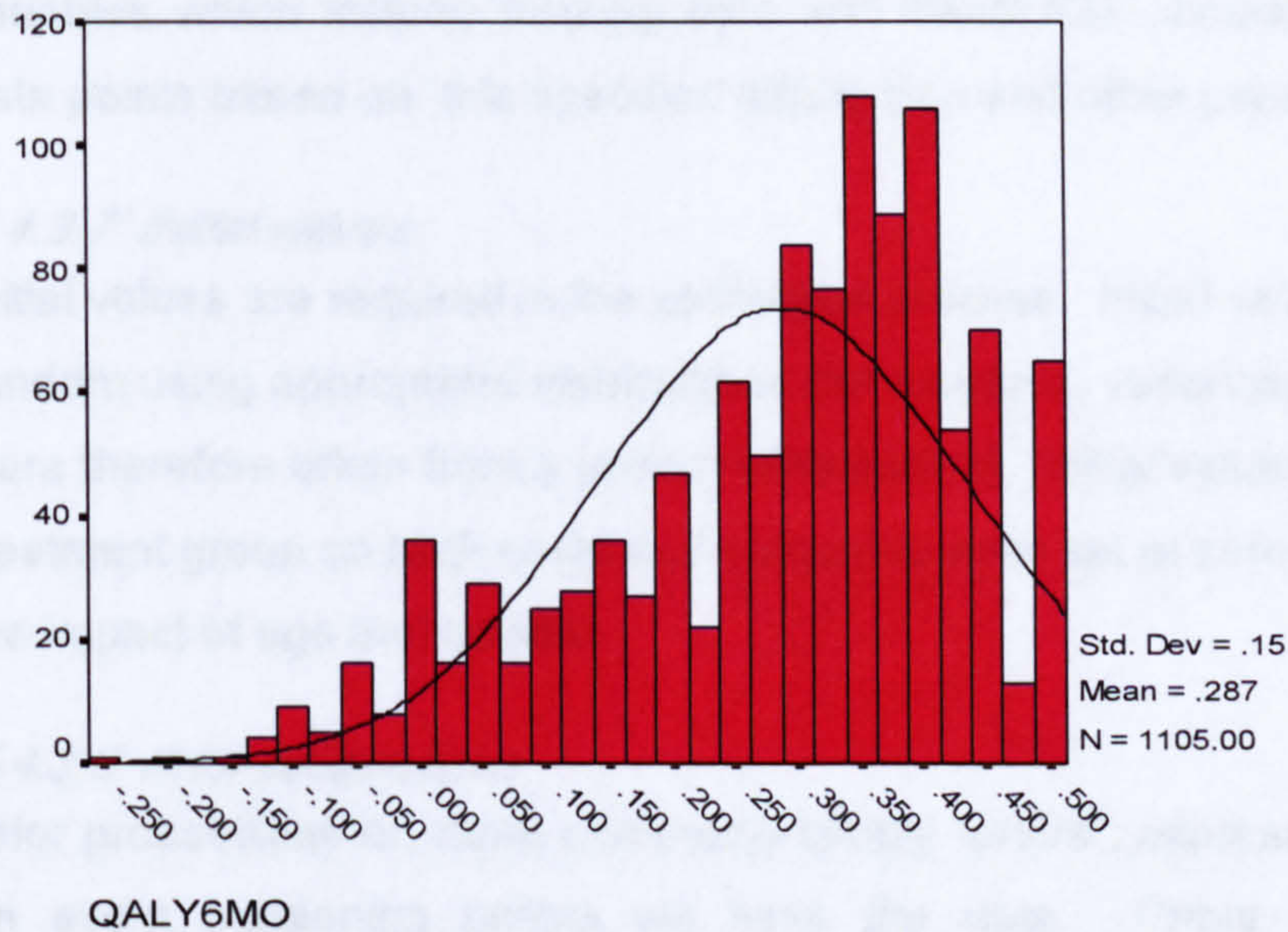


Figure 5.3 Histogram of distribution of QALYs from the 2 datasets combined

5.4.3.5 Correlation between costs and effects

The correlation between costs and effects can be incorporated into the simulation model in WinBUGS with the addition of a further variable in either the cost or QALY simulation. Thus, for example, the cost equation (in a fixed effect framework with study specific EQ-5D as above) could be represented as:

$$C = \alpha + \beta_{C_T}Tx + \beta_{C_S}study + \beta_{int}study*EQ-5D + \beta_{C_X} X_i + QALY + \varepsilon \quad \text{Eq 6}$$

In this example, absolute QALYs are included as a covariate for costs to encapsulate the correlation between costs and effects. Where correlations are large, these covariates are likely to have greater impact. The inclusion of correlation between costs and effects and the impact on model fit and on the conclusions are considered.

5.4.3.6 Missing data

It was assumed for this exercise that there was an ignorable missing data mechanism. Specifically, it was assumed that missing data were “missing at random”, that is that another variable(s) available in the dataset can explain whether

the data are missing or not. This enables the specification of distributions for variables which include missing data and WinBUGS simulates values for missing data points based on this specified distribution and other parameters.

5.4.3.7 Initial values

Initial values are required in the estimation process. Initial values were chosen at random using appropriate distributions (for example, variances must be positive and were therefore taken from a uniform distribution). Initial values for the impact of treatment group on both costs and outcomes were set at zero, as were values for the impact of age and gender.

5.4.3.8 Prior probabilities

Prior probabilities or, more commonly simply “priors”, express the uncertainty about an event happening before we have the data. Priors were assigned to be uninformative (or vague) with zero mean and very high variance. Note that in WinBUGS, normal and log normal distributions require the mean and precision as arguments rather than the mean and variance, thus a uninformative prior would have a very low second argument, for example `dnorm (0,0.0000001)`

5.4.3.9 Number of iterations and burn-in

There is no consensus as to the ideal number of iterations for a MCMC for Bayesian inference. Cooper et al considered that a 5,000 burn-in (where the results of these simulations are discarded) was sufficient for convergence, with their analysis based on a further 15,000 iterations.⁹⁶ Following this example, 20,000 iterations were conducted with the first 5,000 discarded. Checks for convergence and other diagnostics were performed as detailed in the section below.

5.4.3.10 Assessing the model

Models were assessed for convergence using history/trace plots to show that the successive samples move around the modal value. Kernel density was also monitored as non-convergence can manifest itself with a kernel plot that is multi-modal (more than one local maximum)

Autocorrelation, where sequential samples are correlated, can be problematic, and this was also examined using plots.

The models were assessed for their ability to fit the data using the Deviance Information Criterion (DIC) function in WinBUGS which is used to assess model complexity and compare different models.⁹⁷

5.4.4 Sensitivity analyses

5.4.4.1 Intervention costs

The national evaluation of EPP estimated the cost of the intervention as £250. However, Griffiths et al used a less exhaustive methodology and estimated the cost at only £123. This figure is employed as a sensitivity analysis.

5.4.4.2 Initial values

The choice of initial values should not impact on the posterior distribution. However, in order to test this assumption, initial values were varied to assess the impact on results. This is commonly done using more than one “chain” and setting different initial values for each chain. If the chains converge, the impact of different initial values is limited.

5.5 Results

5.5.1 Choice of covariates for the model

5.5.1.1 QALYs

The model potentially contained all the covariates described in section 5.4.3.3. The results of the stepwise regressions are presented below.

Table 5.3 Regression of potential explanatory variables on QAL Ys

	<i>Estimate</i>	<i>Standard Error</i>	<i>p-value</i>
Constant	0.276	0.0098	0.000***
Treatment effect	0.011	0.0048	0.152*
Baseline EQ-5D	0.362	0.0072	0.000***
Study dummy	0.002	0.0055	0.716
Age	0.000	0.000	0.5185
Gender	0.003	0.005	0.601
Self-efficacy baseline	0.001	0.001	0.226
AIC		-2436.3	
****	indicates statistically significant at 0.1% level		
***	indicates statistically significant at 1% level		
**	indicates statistically significant at 5% level		
*	indicates statistically significant at 10% level		

Deleting non-significant variables in a one-by one (stepwise) deletion of the least significant yields the model displayed in Table 5.4. The treatment effect, though not statistically significant, has been retained in the model as it is the variable of interest.

Table 5.4 Stepwise elimination model for QAL Ys

	<i>Estimate</i>	<i>Standard Error</i>	<i>p-value</i>
Constant	0.281	0.0040	0.000***
Treatment effect	0.012	0.0048	0.167*
Baseline EQ-5D	0.362	0.0070	0.000***
Study dummy	0.001	0.0049	0.771
AIC		-2446.5	

The study dummy variable above represents the use of a fixed effects model, where there is a fixed treatment effect on costs and QALYs, but each trial has a study specific intercept. The regression shows that the treatment has a positive effect on QALYs (of about 0.012 QALY) which, predictably lies between the estimates of the two included trials. The only significant covariate is baseline EQ-5D, and this was therefore included in the models used in WinBUGS to determine the appropriate distributional assumptions. Note that the AIC has been reduced in moving from the comprehensive model to the simplified model. This reflects the fact that fewer variables have been included in the regression.

5.5.1.2 Costs

EQ-5D baseline score was a significant predictor of total costs, and when backward stepwise regression was performed, no other potential covariate reached conventional levels of statistical significance in the final model. Though the simple model with baseline EQ-5D only as a covariate had a slightly higher (i.e. worse) AIC score than the model that included all covariates, the differences were small and this simpler model saved considerable time and computational effort in WinBUGS, and was therefore deemed appropriate.

It is asserted therefore that the following specifications include appropriate covariates to model cost and QALY data:

$$C=\alpha + \beta_{C_T}Tx + \beta_{C_S}study + \beta_{C_E}EQ-5D + \epsilon \quad \text{Eq 7}$$

$$Q=\alpha + \beta_{Q_T}Tx + \beta_{Q_S}study + \beta_{Q_E}EQ-5D + \varepsilon \quad \text{Eq 8}$$

Where:

Where Tx is treatment group, SE is self-efficacy
 β_{C_T} and β_{Q_T} is the effect of treatment on costs and QALYs respectively
 β_{C_S} and β_{Q_S} is the effect of study on costs and QALYs respectively
 β_{C_E} and β_{Q_E} is the effect of EQ-5D on costs and QALYs respectively
and ε is the error term

These specifications were then fed into WinBUGS and alternative distributional assumptions were assessed as described in the next section.

5.5.2 Distributional assumptions

The software package WinBUGs was used for the following MCMC simulations to ascertain the most appropriate distributional assumptions.

5.5.2.1 Cost distribution

Section 5.4.3.3 displayed the distribution of costs in the two trials. It was argued that a log-normal distribution for costs may be more appropriate. If this is the case we would expect that the DIC score (see 5.4.3.10) would improve (have a lower value) when we fit a log normal rather than a normal distribution on costs. For illustration therefore, costs with a normal distribution are represented and compared with a log normal distribution.

Table 5.5 Comparison of alternative distributional assumptions on model fit

	Incremental cost (including intervention cost)	Incremental QALY	ICER	DIC*
Normal costs (with no covariates)	207	0.0088	23523	36251
Log normal costs (with no covariates)	262	0.0084	29633	34729

** lower score reflects "better" model fit*

Thus, the DIC indicates that the log normal costs are a better representation of the data than using the normal distribution. This is consistent with examination of the distributions presented earlier.

5.5.2.2 Inclusion of study specific covariates and correlation

Section 5.5.1 presented the results of the regression analysis to identify some important covariates and showed that baseline EQ-5D was an important covariate. The previous section, 5.5.2.1 demonstrated that a log normal distribution for costs was appropriate. Now, we can consider whether the EQ-5D baseline adjustment should be based on a study specific value and whether correlation between costs and effects is important. The results of these analyses are presented in Table 5.6. The first row, for comparison, repeats the final row of Table 5.5. Log normal distributions on costs offer a “better” explanation of the data than normal distributions. In turn, adding a generic EQ-5D covariate and subsequently a study specific EQ-5D covariate show a very slight improvement in the model. However, the inclusion of the correlation between costs and effects actually makes no sizable difference to the model fit.

Table 5.6 Comparison of models including covariates

	Incremental cost (incl. intervention cost)	Incremental QALY	ICER	DIC*
Log normal costs (no covariates)	262	0.0084	29633	34729
Log normal costs with generic EQ- 5D covariate	263	0.0088	29864	34727
Log normal costs with study specific EQ-5D covariate	262	0.0088	29838	34682
Log normal costs with study specific EQ-5D covariate and correlation	261	0.0087	29843	34687

** lower score reflects “better” model fit*

5.5.3 Cost-effectiveness

Table 5.6 also presents the results of the cost-effectiveness analysis in terms of the costs and QALYs generated and the resulting ICER. Each of the log normal cost models gives similar results in terms of costs and QALYs generated and the ICER while the model using a normal distribution for costs result in a slightly lower (better) ICER and higher probability of being cost-effective, but does not fit the data as well. Two log normal cost CEACs are presented as the lower two lines in the graph below, while normal costs are shown as the highest of the lines. This CEAC illustrates the result that using two datasets with different conclusions leads to a

more ambiguous solution. At a cost-effectiveness threshold of £30,000 there is only a 50% probability that the intervention is cost-effective, while at £20,000 per QALY, the probability is only around 20%. There is considerably more uncertainty around the decision using two datasets of IPD, than there was based on the single trial based analysis presented in Chapter 4.

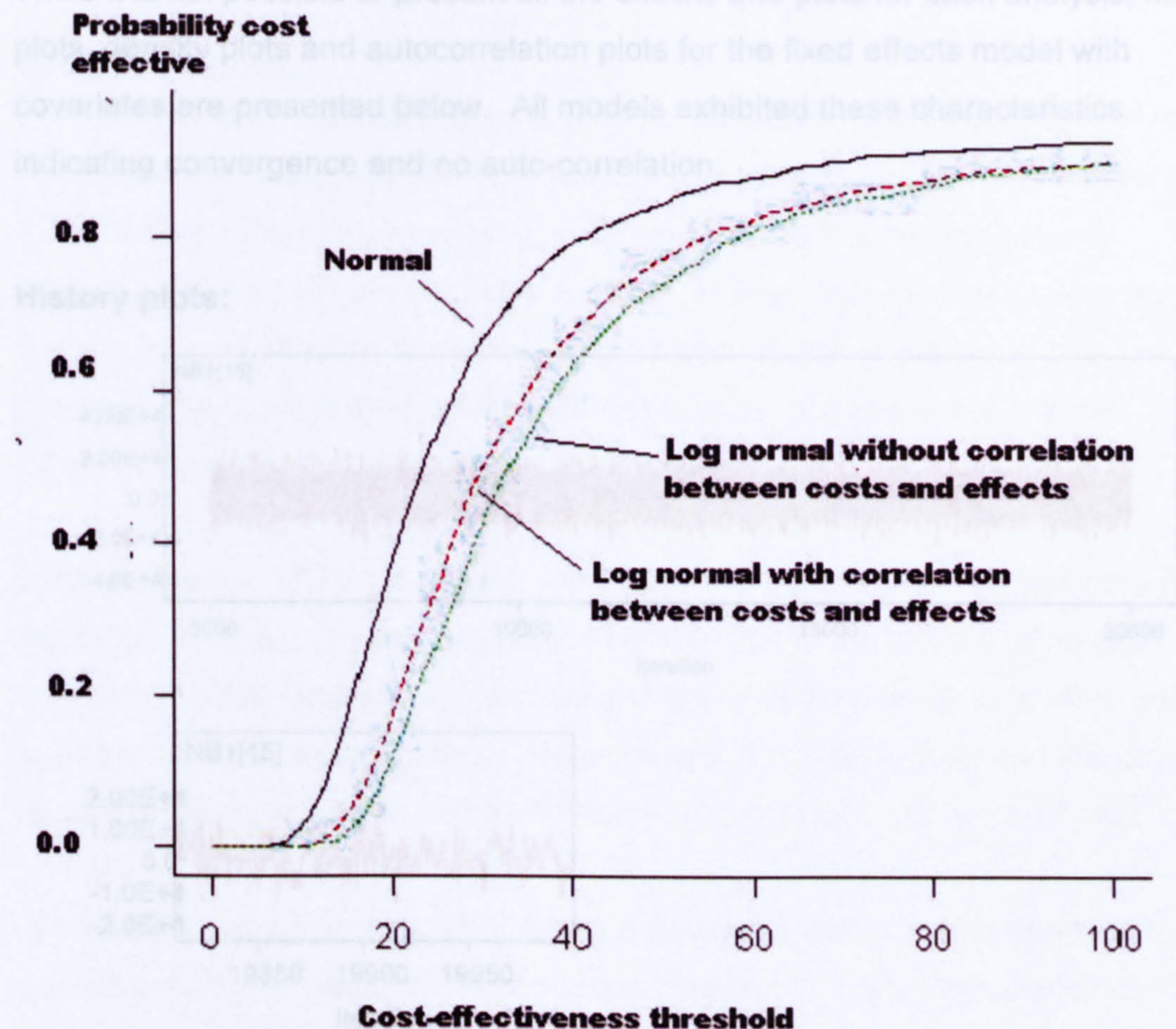


Figure 5.4 Cost-effectiveness acceptability curves with two datasets with alternative distributional assumptions

5.5.4 Sensitivity analysis

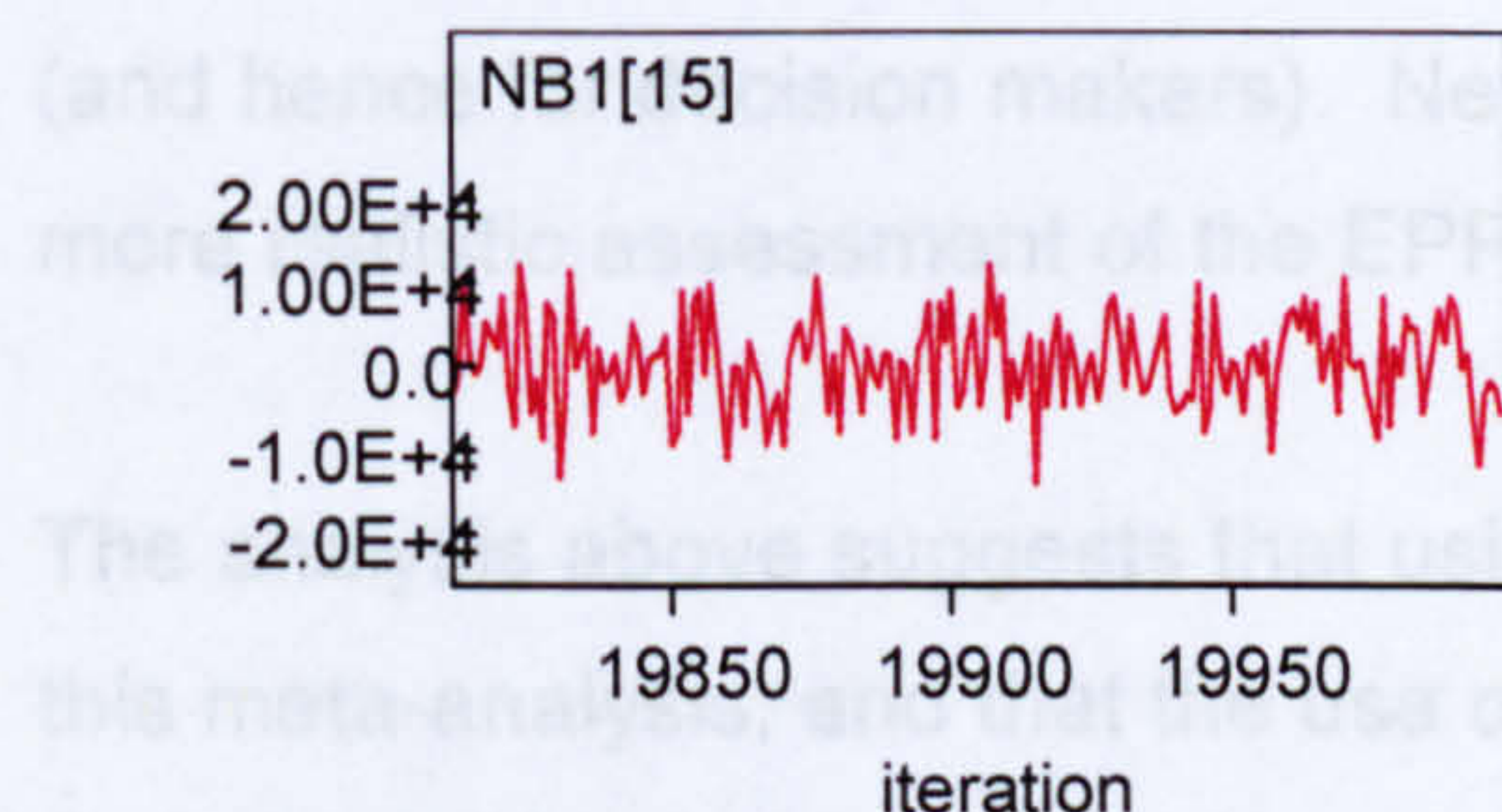
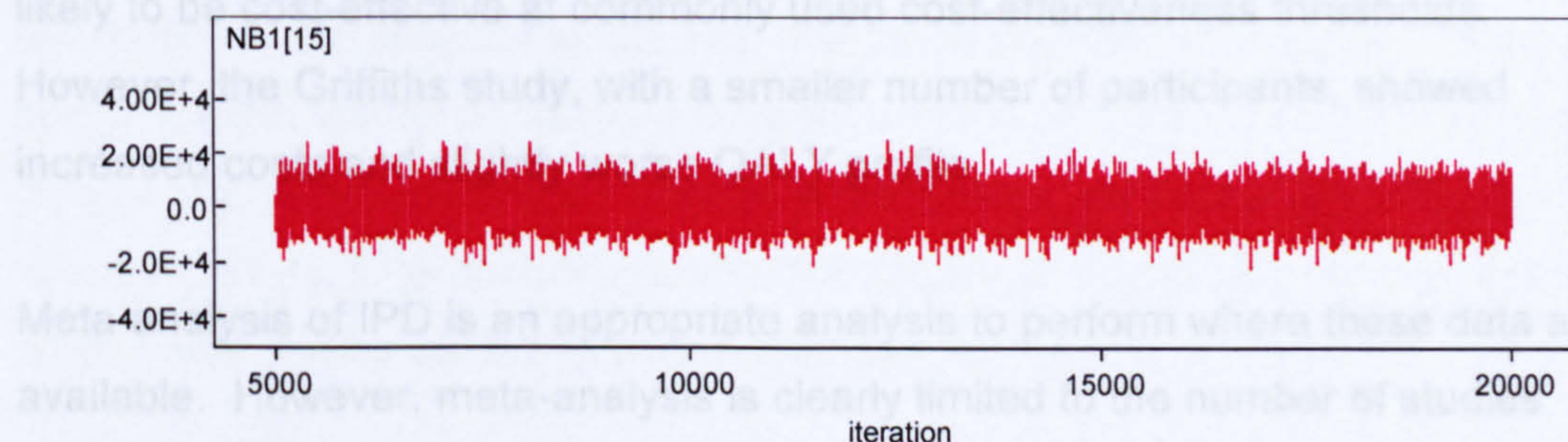
The reduction of the cost of the intervention to £123 as estimated in the Griffiths study, not surprisingly improves the likely cost-effectiveness of the intervention. Using the model that “best” describes the data, that is the model with study specific EQ-5D covariate but no correlation, yields an incremental cost of £185 with the same incremental QALY of 0.0088. This generates an ICER of £21,000 and a

probability of the intervention being cost-effective at £30,000 of around 70% (and around 45% at £20,000 per QALY). The results were not sensitive to the choice of alternative initial values.

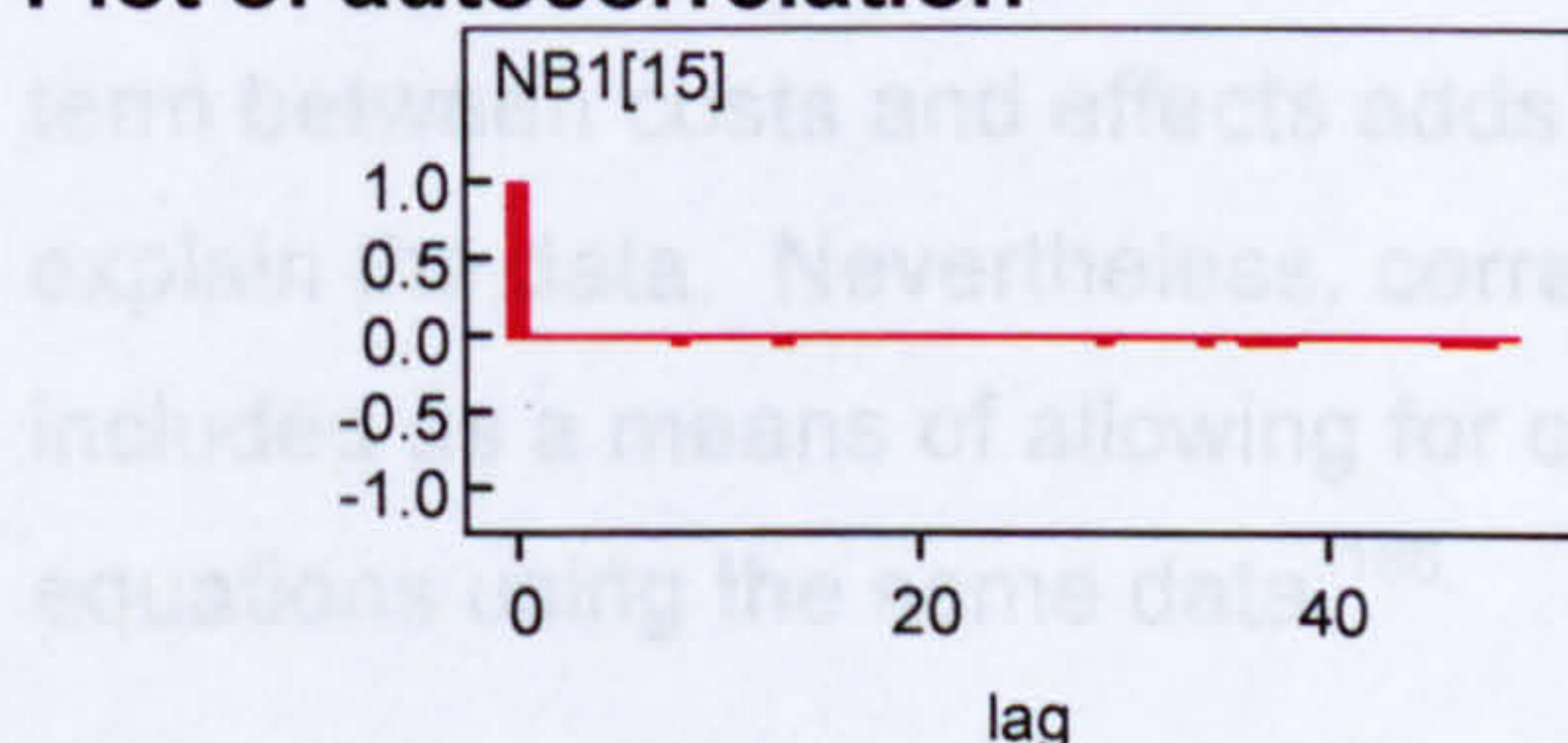
5.5.5 Checks for convergence and auto-correlation

Section 5.4.3.10 described checks that should be carried out on each of the models. While it is not possible to present all the checks and plots for each analysis, history plots, density plots and autocorrelation plots for the fixed effects model with covariates are presented below. All models exhibited these characteristics indicating convergence and no auto-correlation.

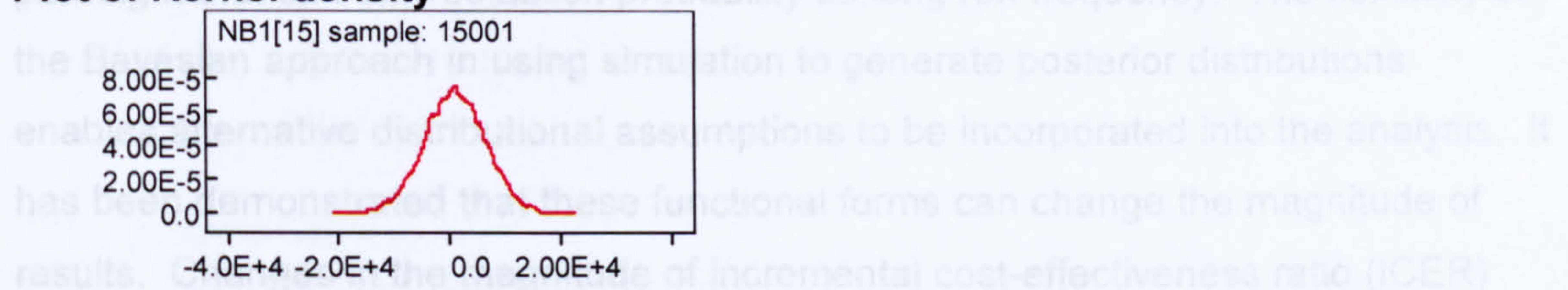
History plots:



Plot of autocorrelation



Plot of Kernel density



5.6 Discussion

The above analysis shows that for a range of assumptions about the type of model, there is considerable uncertainty around the cost-effectiveness of the intervention. This result is not entirely surprising. The larger of the studies, the national evaluation of the EPP based on the CDSMP, showed that the intervention was very likely to be cost-effective at commonly used cost-effectiveness thresholds. However, the Griffiths study, with a smaller number of participants, showed increased costs and slightly worse QALY profile.

Meta-analysis of IPD is an appropriate analysis to perform where these data are available. However, meta-analysis is clearly limited to the number of studies available. In this instance, two studies provide limited evidence for a meta-analysis (and hence for decision makers). Nevertheless, it can be argued that this provides a more realistic assessment of the EPP than the single trial evaluation of the EPP.

The analysis above suggests that using a lognormal distribution is appropriate for this meta-analysis, and that the use of this distribution has a negative impact on the likely cost-effectiveness of the intervention. Including study specific EQ-5D scores improves the model appreciably, though the subsequent addition of a correlation term between costs and effects adds little to the model in terms of the ability to explain the data. Nevertheless, correlation between costs and effects has been included as a means of allowing for correlated error terms between costs and effect equations using the same data.¹⁸⁶

Using Bayesian techniques enables the analyst to assess the impact of employing alternative models on the probability of the intervention being cost-effective. The concept of “probability of cost-effectiveness” is meaningless in a frequentist

paradigm that can only establish probability as long run frequency. The flexibility of the Bayesian approach in using simulation to generate posterior distributions enables alternative distributional assumptions to be incorporated into the analysis. It has been demonstrated that these functional forms can change the magnitude of results. Changes in the magnitude of incremental cost-effectiveness ratio (ICER) can impact on the recommendation to decision makers. In this example, due to uncertainty around the exact threshold value of the QALY, the policy implications are unclear. If the threshold value of a QALY were either £20,000 or £30,000, the recommendation would not alter with the distributional assumptions and the use of covariates. In the former case, the intervention would be deemed *not* cost-effective (as the ICER exceeds £20,000 per QALY), while in the latter the intervention would be deemed cost-effective. However, at a threshold value of £25,000, changes in distributional assumptions would alter the conclusions of the analysis. Using the normal distribution for costs the intervention would be deemed cost-effective, whilst using a log normal distribution the intervention would not be considered cost-effective. Thus, in terms of implications for policy, the distributional assumptions employed in the model are of critical importance.

5.7 Conclusion

The meta-analysis described above demonstrates the importance of employing the appropriate distributional assumptions and covariates. While meta-analysis of IPD may be appropriate in this instance, it is also limited in that there may be more data available on the effectiveness of this intervention, which could be used to give a better estimate of the uncertainty around the adoption decision. The identification of this evidence and their subsequent synthesis with existing data is considered in the next chapter.

Chapter 6. Evidence synthesis of IPD and summary statistics of surrogate outcomes

6.1 Introduction

The previous chapter extended the single trial based analysis to incorporate evidence from another UK based randomised controlled trial. However, in this instance, the inclusion of this additional evidence was unlikely to reduce the uncertainty around the decision over the adoption of the intervention. Indeed, at commonly quoted threshold values of a QALY, the uncertainty around the adoption decision was markedly increased. These two trials present the only evidence from UK based RCTs of the CDSMP intervention.

However, we know from Chapter 3, that there are other studies outside the UK that may provide relevant evidence to inform the adoption decision. As it is frequently stated that economic evaluation should incorporate all relevant evidence,^{38, 62, 84, 187, 188} it is reasonable to consider their inclusion in the analysis to inform the decision. Indeed, decision makers in the UK¹² and other jurisdictions¹¹ provide guidance advising the desirability of the inclusions of *a//*relevant data.

There is not, however, a simple dichotomy between relevant and irrelevant evidence. “Relevance” is context specific and involves a value judgement, ideally based on some scientific criteria. However, in conducting evidence syntheses, it is necessary to define what is included and what is excluded. Relevant evidence may vary not only across geographical regions and different conditions, but also according to who is making the decision on behalf of whom.

At one extreme, in the case of the CDSMP, the single trial based analysis reported in Chapter 4 could be considered to contain the only relevant information, as this was the only trial of the CDSMP that was based on a representative sample of the UK public.^{66, 162} However, one could also argue that another trial of the CDSMP in the UK,⁸³ albeit in a minority population, could also be considered relevant. The evidence base for the decision is therefore extended in Chapter 5 by synthesising the two trials.^{66, 83, 162}

We could go a step further. Studies showing the effectiveness of the CDSMP in other countries, in different populations, in various conditions, over different time periods using alternative outcome measures with different trial designs could be of interest. The extent to which these studies are relevant is open to discussion. We know from Chapter 3 that there are studies that estimate the effectiveness of the CDSMP, but using non-generic outcome measures. However, these studies, and any others that may be identified by the search strategy (described later), are *potentially* relevant. How they might be included, the assumptions that this necessitates, and the impact on the results, conclusions and recommendations of the analysis are the focus of this chapter.

The previously identified studies have different outcome measures from the generic one used in Chapters 4 and 5. These outcome measures can be considered to be surrogate (or intermediate) outcomes that are linked to health and are affected by the intervention (CDSMP). By treating these outcomes as surrogates and establishing the link between the surrogate and final outcome (in this instance health, as measured by the QALY), we can incorporate this evidence into the decision problem.

As well as the use of surrogate outcomes, individual patient data are not available for inclusion in the analysis. Synthesising these different data types (IPD from the two trials used in Chapter 5 with aggregate data from the published literature) requires the use of additional modelling techniques and concepts and these are now described.

The objectives of this chapter are three-fold. Firstly, to develop a framework for synthesising published (aggregate) data with the trial based data presented in the previous chapter. Secondly, to identify any other potentially relevant studies that may fit into this model and improve the analysis of the decision problem. Finally, to examine the impact that choice of evidence and what constitutes relevance may have on the adoption decision and the decision to conduct further research. A version of this chapter has been presented at Health Economics Study Group.¹⁸⁹

6.2 Modelling concepts

6.2.1 *Synthesising IPD and AD*

Chapter 5 described the IPD trial data available for inclusion in the model. However, the review in Chapter 3 indicated that there were other published data examining the effectiveness of the CDSMP. These data were available as aggregate data (AD) rather than IPD, therefore, to include these data requires some form of synthesis of IPD and AD. While there is some guidance in the literature as to the preferred conduct of an IPD meta-analysis,^{181, 190} surprisingly, there is little literature around combining AD and IPD.

A similar scenario occurs in IPD meta-analysis which has been described as the gold standard of evidence synthesis.¹⁹¹ In IPD meta-analysis, any studies that do not provide IPD essentially require a decision about how to incorporate AD within the IPD meta-analysis. The usual approach in IPD evidence synthesis has been to exclude these studies where “sufficient” IPD are available and to consider the impact that omitted results may have in a sensitivity analysis.^{192, 193} Clearly, both of these approaches do not make use of all the available evidence and also it is not clear how analysts should proceed when the level of IPD is insufficient (or, indeed, what level constitutes “sufficiency”). While this is unlikely to be an issue in specialties where there is a well developed network with an established history of synthesising evidence, such as in the cancer field, in other areas where these networks are not developed this may provide a sizable problem.

The only paper identified that explicitly examined the synthesis of IPD and AD is the paper by Sutton et al, where the authors consider a meta-analysis where 12/37 trials provided usable IPD with 25/37 only having AD available.¹⁹⁴ An additional complication of this study was that a sub group of both the IPD available studies and the AD available studies were clustered trials. In essence the authors used a meta-analysis model that allows the incorporation of evidence from different study designs. We used a similar methodology (described below) though there were differences due to the nature of the data (our data were not binary outcome data) and the fact that the outcome used in the synthesis below was a surrogate outcome measure. The issue of surrogacy is discussed in the following section.

6.2.2 The use of surrogate outcomes and the assumption of conditional independence

Surrogate outcomes have been defined as “outcome measures that are not of direct practical importance but are believed to reflect outcomes that are important”.¹⁹⁵ For example, we may build a model for drug A in treatment of heart disease, but have only data of the effectiveness of drug A in lowering levels of cholesterol or hypertension. In this instance, cholesterol level or hypertension is a surrogate for the final outcome measure (which may be life expectancy or QALYs, for example). It is assumed that these surrogates would impact on final outcome. Perhaps more importantly it is assumed that *only* these surrogate outcomes impact on final outcome. This requirement is the assumption of *conditional independence*, such that once we know the value of a surrogate, the treatment effect no longer has any impact on outcomes. Formally,

$$P(A|C) = P(A|B,C)$$

Where A is (for example) QALYs, B is treatment group and C is level of autonomy. The knowledge of B (treatment group) imparts no further knowledge of the value of A, given that we know the level of autonomy.

However, there are examples from the surrogate literature where there are direct treatment effects that have been ignored and the surrogate has subsequently been discredited.¹⁹⁶ A frequently quoted example is that of CD4+ counts in HIV trials, where CD4+ count may be improved by the introduction of certain drugs. While CD4+ count is a good measure of survival, and the drugs have a positive impact on CD4+ count, other adverse effects of the drugs may actually reduce survival.¹⁹⁷ This example demonstrates the potential dangers of assuming conditional independence, as not all of the important impact on final outcomes was picked up via knowing the CD4+ status. Nevertheless, the use of surrogates in the economic evaluation literature is common as shown in the review in Chapter 3, where the majority of studies used an outcome measure that (it was hypothesised) acted as a surrogate for the final outcome of interest (for example, health).

There is a considerable literature on the subject of surrogate outcomes. Much of this literature is concerned with hypothesis testing rather than on estimation, which is the primary aim of this chapter. However, recent reviews of the literature have identified a range of methods of testing for the suitability of surrogates as a “true” marker for final endpoints.^{198, 199} Unfortunately, these methods, whilst yielding a numerical estimate of the validity of markers as surrogates, have no guidance as to what figure is required to achieve validity, or whether there is an absolute level that should be met across all conditions. It could be argued that complex interventions, where the “active ingredient(s)” is/are less easy to specify,¹⁶⁷ may have more mechanisms impacting on final outcome than simple interventions (such as a drug for asthma). Does this suggest (or allow) that the level of validity of a single surrogate amongst many could or should have a lower threshold for acceptability than one where we are more certain of the (biological) mechanism? This reduces to a question about our beliefs about the structure of the model and whether it is correct. If we believe the model is structured appropriately, and that everything goes through the single surrogate, then the fact that it is a “poor” surrogate may not be so important, we just might not be very good at measuring it. If however, we believe that the reason it is a “poor” surrogate is due to the exclusion of other important covariates or mechanisms, then we are unsure of the structure of the model and may require more evidence that the structure is appropriate.

There are a number of measures of the adequacy of surrogates. One of the most commonly used measures to assess the suitability of a surrogate is the “proportion of treatment effect explained” (PTE) method, where PTE is defined as:

$$\text{PTE} = 1 - (\tau_{\text{adjusted}} / \tau_{\text{unadjusted}})$$

Where τ_{adjusted} is the treatment effect adjusted for the surrogate and $\tau_{\text{unadjusted}}$ is the treatment effect unadjusted for the surrogate

As stated previously, there is little guidance as to what might be an acceptable figure in this example, though previous authors have suggested that values “close to” zero would indicate an invalid surrogate,^{200, 201} while a value of 1 indicates a perfect surrogate.¹⁹⁹ However, PTE has been criticised on the grounds that it is not

bounded by zero and 1 and is therefore difficult to interpret.¹⁹⁹ This is a developing area in biostatistics, but at present there is no consensus as to the appropriate method for validating a surrogate outcome. However, for the purposes of this analysis we are more interested in estimation as oppose to hypothesis testing. That is, we wish to estimate (or predict) a value with a suitable measure of uncertainty, rather than test whether a certain value is likely or not (subject to some pre-determined level). It was argued in Chapter 2, that hypothesis testing is not particularly suited for decision making. In this case, there may be additional variables that aid estimation or prediction; ideally these should not be excluded simply due to their lack of statistical significance.

6.2.3 Exchangeability

The concept of exchangeability in economic evaluation encompasses both the likely exchangeability of both the relative treatment effect on costs and the relative treatment effect on outcome data. The assumption of exchangeability between two trials requires that there are no *a priori* reasons for expecting a systematically different (higher or lower) estimate of the relative treatment effect on costs and/or effects between the two trials.^{202, 203}

As with relevance, exchangeability can be represented on a continuum ranging from “not exchangeable” to “totally exchangeable” via “partially exchangeable”.²⁰⁴

Exchangeability, again like relevance, will depend on the context. For example, a UK decision maker may decide that evidence from both the UK based RCTs is relevant and exchangeable. However a decision maker with responsibility for rolling out the EPP in a Bangladeshi population in London, may feel that the national evaluation (based across the whole of the UK with few minorities represented) has less relevance and that the treatment effect on costs and QALYs are not exchangeable. Extending this argument, it should be apparent that there is no clear dichotomy between exchangeable and not exchangeable, that is, there may be degrees of exchangeability, so that some evidence is considered to be more exchangeable than other evidence. In particular there is often a distinction made between the exchangeability of the treatment effect on cost data and the exchangeability of the treatment effect on outcomes data.²⁰⁵

6.2.3.1 Exchangeability of costs vs exchangeability of outcomes

It is plausible that the exchangeability of the relative treatment effect on costs may differ from the exchangeability of relative treatment effect on outcome evidence. It has been assumed in some analyses that for cost data this is not exchangeable,²⁰⁶⁻²⁰⁸ while the relative treatment effect on outcomes is exchangeable. It is necessary to be explicit when setting out the decision problem and address issues of relevance and exchangeability at the outset. For the purposes of this analysis, both datasets of IPD are UK based using a common unit cost base and a common perspective and adjusted for differential timing. In this instance, and for these data, it is assumed that both the relative treatment effect on cost data and outcome data are exchangeable.

Intuitively it is plausible that the relative treatment effects on cost and outcome data are at least *as* exchangeable as the relative treatment effects on self-efficacy, as the relative treatment effects on self-efficacy are measured in different populations, in different settings. For the purposes of the analysis below, it is assumed that the treatment effect on cost data from the two UK trials is exchangeable, and the relative treatment effect on outcomes from the UK and elsewhere are exchangeable.

6.2.3.2 Exchangeability of outcome measures at differing time points

Studies supplying aggregate data only, had differential follow-up periods. An earlier chapter referred to the adjustment made to QALYs in one trial (with IPD) to ensure compatibility with the other trial. This is a requirement for QALYs as by definition QALYs are time dependent (being a product of utility score associated with a state and the time spent in that state). It is not clear *a priori* whether self-efficacy is time dependent, or if it is time dependent, what the nature of that relationship is. The relationship is not simply one that increases/decreases monotonically with time as some studies with shorter follow-up demonstrate greater changes in self-efficacy. If data were available that considered self-efficacy scores at different time points, it may be possible to model the impact of time on self-efficacy. Unfortunately, such data are not available, and differences in the levels of self-efficacy across studies are equally likely to be a result of different characteristics of the study population,

the study design or other factor. Therefore, for the purposes of this analysis, an assumption is made that self-efficacy is time independent.

6.2.4 Fixed and random effect models

As described in Chapter 5, meta-analysis of IPD should use a trial identifier in the model. This allows differences between trials to be included and is more realistic than simple pooling of the data without this identifier (effectively treating the data as though it were from one large “mega-trial”). The analysis is then a “stratified” analysis (stratified by trial) and we have the choice between conducting a fixed effect or random effect model. With the fixed effect model, each trial within the model has a different baseline level of costs and QALYs. The treatment effect (the effect of CDSMP on either costs or QALYs) is considered common to all trials but each trial has a specific constant (equivalently the estimated regression line for each trial is raised/lowered by a specific amount), for both costs and QALYs. In contrast, a random effects model assumes that differences in the relative treatment effects on either costs or QALYs between trials are drawn from a distribution of random disturbances.

Again, there is no consensus as to whether fixed effects or random effects are preferable. In this instance we have a small number of studies for the meta-analysis. With such a small number, creation of a distribution around these estimates is impractical, and for the purposes of this analysis therefore, a fixed effects model was chosen.

6.2.5 Uncertainty in economic evaluation

Cost-effectiveness acceptability curves (CEACs) have been described in an earlier chapter. Briefly CEACs display in graphical format the probability that an intervention is cost-effective for a given threshold value of a QALY. In practise, there is always some level of uncertainty around the cost-effectiveness of an intervention, so that there is some possibility of mistakenly considering an intervention cost-effective at a particular threshold, when in reality it is not cost-effective at that threshold. This is due to the uncertainty around the input parameter estimates which are explicitly included in probabilistic sensitivity analysis, the

currently recommended technique⁸⁹ to “reflect the combined implications of uncertainty in parameters”.¹² If we have perfect information, equivalent to there being no parameter uncertainty, the correct decision will always be made. As stated above, this will not happen in practise and expected value of perfect information (EVPI) is a methodology that enables us to place a monetary value on the consequences of making a wrong decision (at various threshold values of a QALY) and represents the maximum value of further research that would eliminate uncertainty from the decision to be made, *given the current level of evidence*.²⁰⁹ As with CEACs, the EVPI can be plotted against threshold values of the QALY and provides a necessary but not sufficient condition for the conduct of future research (the EVPI must exceed the cost of conducting the research). The analysis below assesses the impact of different model structures on the likely cost-effectiveness of the CDSMP intervention, but also uses the EVPI to assess how alternative structural assumptions may yield different estimates of the value of conducting further research.

6.3 Model development

Section 6.2 described the modelling concepts required to allow using both IPD and AD in a model. These methods enable the use of alternative data types that are available for this analysis. In summary we have six potential types of data for use in the analysis, namely, IPD linking treatment and costs/QALY directly, IPD linking treatment to self-efficacy, IPD linking self-efficacy and QALY, or aggregate data (AD) for any of these relationships.

The analysis below allows IPD and AD from different sources with different outcome measures. A graphical representation of the analysis is presented below in Figure 6.1. While this chapter describes a specific example of synthesising IPD of final outcome and surrogate outcome AD, the model and the code are transferable to any of the following:

- i) synthesis of IPD surrogate and IPD final outcome
- ii) synthesis of AD surrogate and AD final outcome

iii) synthesis of IPD surrogate and AD final outcome

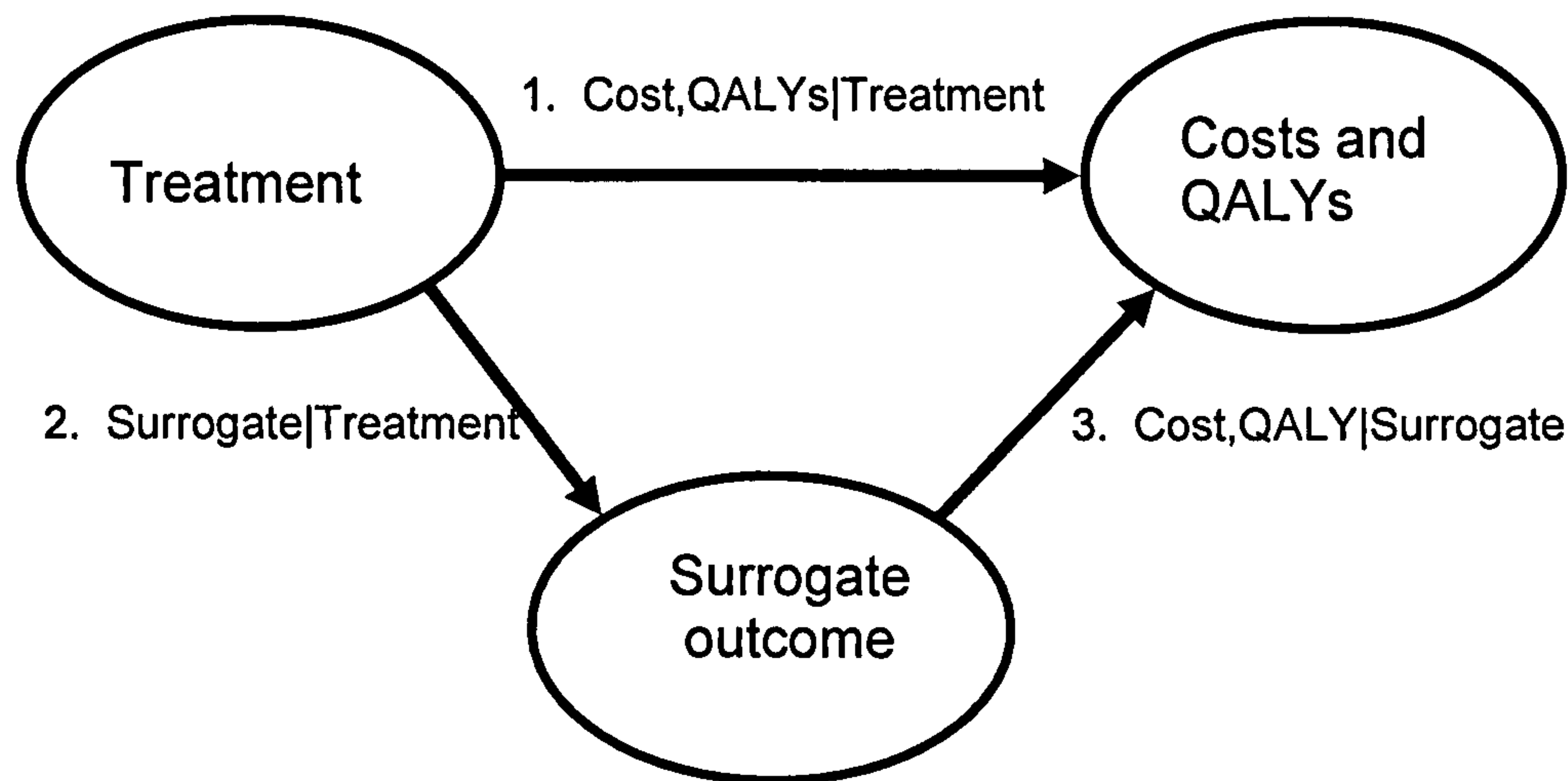


Figure 6.1 Potential mechanisms for CDSMP to impact on costs/QALYs

In this representation, there are two mutually exclusive ways that treatment may impact on costs and QALYs and one that is a hybrid of these two. The direct method, which is common to most trial based evaluations (labelled 1 above), simply estimates the impact of treatment effect on outcomes (costs and QALYs). The indirect method considers the impact of treatment on surrogate outcome and then the impact of surrogate outcome on final outcome (this is the product of 2 and 3 above). The “mixed” method combines both the direct and indirect model to incorporate both direct evidence of treatment effect on final outcomes and the indirect effect which works via the surrogate.

In the specific example considered in this chapter, the CDSMP is the treatment and self-efficacy is the surrogate outcome. Using the graphical representation above, the IPD trials described in Chapter 5 provided *direct* evidence of the impact of the CDSMP intervention on QALYs (labelled 1 above). However, these trials also measured treatment effect on self-efficacy and the effect of self-efficacy on final outcome (QALY) and therefore they also contribute to the estimates 2 and 3 above. Thus these trials provide *indirect* evidence via self-efficacy (that is treatment group impact on self-efficacy which impacts on QALY).

We know from the review in Chapter 3 that there are other sources of evidence that use self-efficacy as an outcome measure. These studies, and any others like them that are identified in the search below, inform the link between treatment (CDSMP) and surrogate outcome (self-efficacy) which is labelled 2 in the representation above. These studies did not include QALYs as an outcome measure and can therefore only contribute *indirectly* via the effect of CDSMP on self-efficacy (relationship 2 in the above graph). Thus, in this instance, they do not contribute to the direct evidence, though in principle at least, such data could be incorporated subject to satisfying certain assumptions (if we had AD for a study using CDSMP and measuring QALYs and costs).

In addition, though in this example it is only the IPD that contribute to relationship 3 (between self-efficacy and QALYs), this is due to a lack of identifiable evidence from other sources. The assumptions associated with each model are described in more detail below.

6.3.1 The “direct” model.

The direct trial based analysis estimates the effect of treatment group on final outcomes, costs and QALYs. Thus in a regression framework, treatment group would be used as an explanatory variable for the outcomes of interest (costs and QALYs). The model developed in Chapter 5 is considered appropriate.

This model allows evidence to be incorporated *only* where the treatment is CDSMP and the outcomes measured are costs and QALYs. In this example, the only direct evidence we have linking treatment (CDSMP) to costs and QALYs is the IPD from two clinical trials.^{66, 83} No indirect evidence via self-efficacy is admissible in this model, but the model does not require additional conditional dependence assumptions to be satisfied.

For ease of analysis, and as there were only two sets of data available (only IPD data informed the direct relationship), a fixed effect rather than a random effect model was used, see above. The assumptions in this model are that the cost and QALY data from both of these trials are completely independent (in that if we

observe costs (QALYs) in one trial, the costs (QALYs) in the second trial are the same as if we had not observed the first). However, an additional assumption is that the costs and QALYs from these trials are exchangeable. As both trials for which IPD were available were UK based and a common costing framework was employed, this is deemed a reasonable assumption for costs. The exchangeability of the relative treatment effect between two UK based trials is less controversial and is a commonly accepted methodology. However, a degree of caution should be exercised in this instance as the data from one of the trials was in a specific ethnic population which may reduce the degree of exchangeability of the relative treatment effect. A graphical representation of this model is presented below. The relationship between costs/QALYs and the treatment is represented by the expression over the arrow, Costs,QALYs|Tx (Costs and or QALYs, conditional on treatment group).

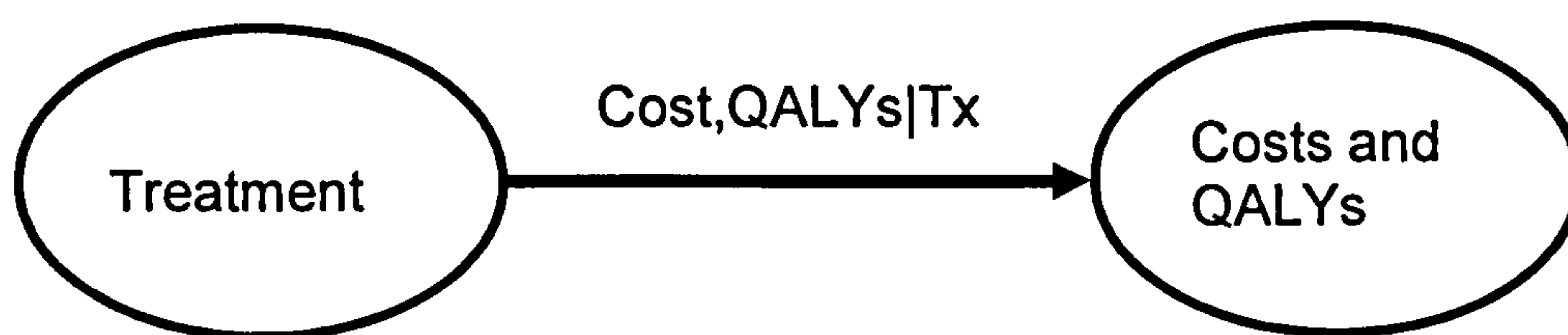


Figure 6.2 The “direct” model

6.3.2 The “indirect” model

In this model, we consider an indirect analysis using effect of treatment group on self-efficacy and then self-efficacy on QALY as the mechanism, that is, we consider self-efficacy as a surrogate for the final outcome measure which in this instance is cost or QALY. No direct mechanism is considered in this model. In essence, this model is similar to many decision analytic models in the health field where there are no data linking treatment and final outcome directly, merely an intermediate or surrogate outcome that, it is asserted, “predicts” final outcome. As described in section 6.2.2, a requirement of these models is that of conditional independence. In this example, we can state that given the value of self-efficacy, the measure of treatment effect is independent of treatment choice. Equivalently, we can say that the treatment has no impact on costs and QALYs, *other than the effect* via self-efficacy.

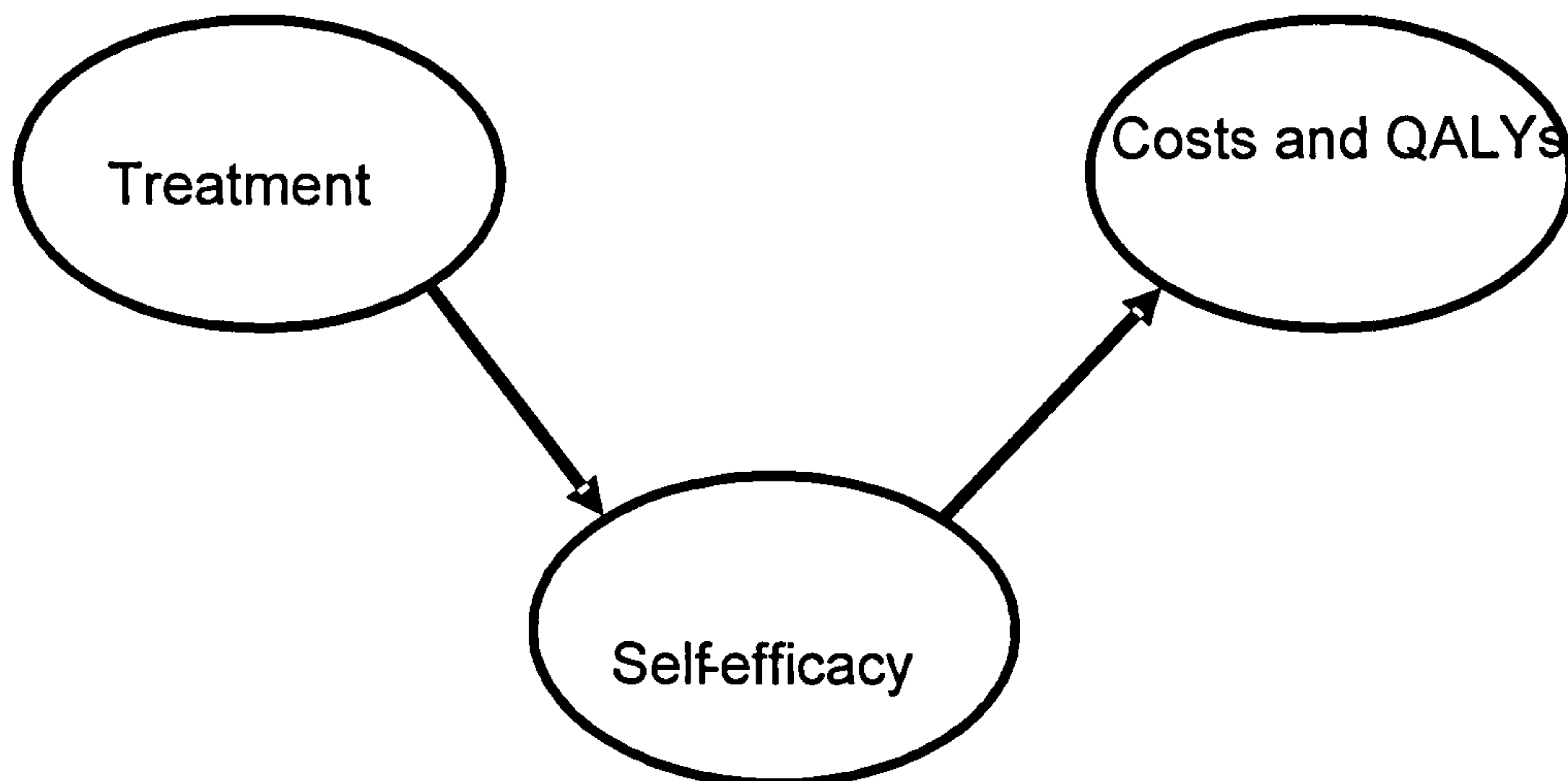


Figure 6.3 The “indirect” model

In addition to this strong assumption is that of exchangeability of relative treatment effect on outcomes. That is, the impact of the CDSMP intervention on self-efficacy is similar in patient groups with differing characteristics in different settings. These trials were conducted in different countries (US, UK, Australia, China) and the assumption is that these trials supply exchangeable estimates. While it is commonly assumed that treatment effects on outcomes are transferable,²¹⁰ this is largely in conditions where there is a “hard” outcome measure such as mortality in heart disease. It is less clear whether a “psychological” outcome such as self-efficacy (which is defined as confidence in ability to manage condition) is as robust to international and inter-ethnic exchangeability as harder measures of outcome. The direct analysis (in section 6.3.1 above) also assumes a common relative treatment effect.

6.3.3 The “mixed” model

The third model again extends the direct and indirect models to include both direct and indirect mechanisms. That is, the direct effect of treatment group on QALYs is considered in the analysis and the indirect analysis where QALYs are affected via the impact of treatment group on self-efficacy and the impact of self-efficacy on QALYs.

Again, both IPD and AD can be used to inform this model in principle. In this example, IPD informs the direct relationship as there were no other sources linking CDSMP treatment directly to QALYs, as well as the relationship between self-efficacy and QALYs for the same reason. IPD and AD inform the relationship between CDSMP treatment and self-efficacy.

Clearly, this model is a hybrid of the previous models in that it uses both direct and indirect mechanisms. A consequence of this is that the mixed model requires both the assumptions associated with both the direct and indirect models

6.3.4 Model estimates

The three models, with their assumptions are presented in Table 6.1. This table also shows the coefficients we are interested in for estimation purposes and how the model estimates costs and QALYs.

The table describes the three models in regression terms. In the direct model, we estimate the impact of treatment group on costs (QALYs) with a study specific constant, as we have used a fixed effect model indicating a common treatment effect on costs but a different baseline value for each study. An error term is also included to capture the effects of any omitted variable. The coefficient of interest, the effect of the treatment on costs (or QALYs) is represented by β_{CT} (or β_{QT}). This model used only the IPD available as no other data directly linking CDSMP and costs (QALYs) were identified.

The indirect model requires estimates of the effect of treatment group on self-efficacy and subsequently, the impact of self-efficacy on costs (QALYs). The treatment group effect on self-efficacy is represented by β_{ST} , while the effect of self-efficacy on costs (QALYs) is represented by β_{CS} (β_{QS}). Thus if, the effect of treatment group on self-efficacy and then self-efficacy on costs (QALYs) is large, the impact of treatment group on costs (QALYs) via self-efficacy (the indirect estimate) will also be high. This is represented by the product of the two coefficients, that is, $\beta_{CS} \cdot \beta_{ST}$ (for costs) or $\beta_{QS} \cdot \beta_{ST}$ (for QALYs).

The mixed model combined both direct and indirect evidence. Not surprisingly, the estimate of interest is the sum of the direct and indirect models, that is, $\beta_{CS}.\beta_{ST} + \beta_{CT}$ for costs and $\beta_{QS}.\beta_{ST} + \beta_{QT}$ for QALYs. The first term in these expressions $\beta_{CS}.\beta_{ST}$ (or $\beta_{QS}.\beta_{ST}$) is the indirect via self-efficacy of treatment group on costs (or QALYs), while the second term β_{CT} (or β_{QT}) shows the direct relationship of treatment group on costs (or QALYs).

6.3.5 Summary of the models

In summary, three models were developed. The models use different types of data from different sources using different endpoints. All models incorporate evidence that is potentially relevant. While it is important that economic evaluations incorporate “relevant” evidence, there is not a simple dichotomy between relevant and irrelevant evidence. The choice of model (direct, indirect, mixed) will influence what constitutes relevant evidence and alternative structures allow the incorporation of different types of evidence. There is a small but growing literature on the subject of structural uncertainty in the use of decision analytic models with suggestions as to how this may be addressed.^{67, 211, 212} Of course, the model used is also dependent on the evidence available. It is accepted that if there is no evidence linking treatment to costs and QALYs *directly*, then other structures have to be employed. However, analysts should be aware of the assumptions made within these models and the potentially different estimates of cost-effectiveness and valuation of future research that these different analytical approaches can yield.

6.4 Methods

6.4.1 Model assumption

One aim of Chapter 5 was to assess potential covariates and distributional assumptions. The conclusions of the chapter were that lognormally distributed costs were appropriate, while study specific EQ-5D should be used as a covariate. The addition of a term capturing the correlation between costs and effects had little impact on the “fit” of the model, but was deemed appropriate as previous authors have stressed the importance of including this correlation.¹⁸⁶

Thus the base case model assumed that the cost data were lognormally distributed and used baseline (study specific) EQ-5D as a covariate where this was available. Correlation between costs and effects was also included in the base case analysis. Any study that provided evidence of the effectiveness of CDSMP was included in the base case to address the decision problem more fully by informing the relationship 2 in figure 6.1 (the effect of treatment on self-efficacy). Studies were identified using methods described in 6.4.2 below.

6.4.2 Search methods

A literature search was conducted (see Appendix D for search strategy) to identify all trials that may provide data linking the treatment (CDSMP) to the (hypothesised) surrogate outcome, self-efficacy. Briefly, any trial that reported self care, self help, self-efficacy or similar term was targeted in the search strategy. Thus any paper informing the relationship labelled 2 in the representation below should have been identified. In addition, studies that may inform the third relationship (between self-efficacy and QALYs were also searched). This strategy is presented in Appendix D.

6.4.3 Software

As in the previous chapters, this analysis takes a Bayesian perspective, with parameters treated as random variables to generate posterior distributions for these parameters based on the model or estimation procedure, the choice of initial values and the priors assigned. As asserted previously, Bayesian techniques using commonly available software are relatively easy to conduct and allow a range of distributional assumptions to be incorporated into the model, whilst retaining an intuitive meaning that is useful for decision making.

Table 6.1 Direct, indirect and mixed models

	Estimation	Prediction	Key Assumption(s)	Data employed in this analysis
Direct	$C = \alpha + \beta_{CT}Tx + \epsilon$ $Q = \alpha + \beta_{QT}Tx + \epsilon$	$\Delta C = \beta_{CT}$ $\Delta Q = \beta_{QT}$	Fixed effects model for synthesis of IPD assumes that treatment effect is average of the two trials.	IPD only
Indirect	$C = \alpha + \beta_{CS}SE + \epsilon$ $Q = \alpha + \beta_{QS}SE + \epsilon$ $SE = \alpha + \beta_{ST}Tx + \epsilon$	$\Delta C = \beta_{CS}\beta_{ST}$ $\Delta Q = \beta_{QS}\beta_{ST}$	Exchangeability of cost and QALY data between UK IPD trials Assumes that effect of treatment on costs and QALYs in trials is transferable to general population Conditional independence required. Treatment ONLY works through the surrogate endpoint self-efficacy	IPD and AD
Mixed	$C = \alpha + \beta_{CS}SE + \beta_{CT}Tx + \epsilon$ $Q = \alpha + \beta_{QS}SE + \beta_{QT}Tx + \epsilon$ $SE = \alpha + \beta_{ST}Tx + \epsilon$	$\Delta C = \beta_{CS}\beta_{ST} + \beta_{CT}$ $\Delta Q = \beta_{QS}\beta_{ST} + \beta_{QT}$	Exchangeability of relationships between treatment and self-efficacy and self-efficacy and cost and QALY data between trials Assumes that effect of treatment on self-efficacy and self-efficacy on costs and QALYs in trials is transferable to general population Combination of both of above.	IPD and AD

Where Tx is treatment group, SE is self-efficacy
 β_{CT} is the effect of treatment (CDSMP) on costs
 β_{QT} is the effect of treatment (CDSMP) on QALYs
 β_{CS} is the effect of self-efficacy on costs
 β_{QS} is the effect of self-efficacy on QALYs
 β_{ST} is the effect of treatment (CDSMP) on self-efficacy

Each of the models described below has been run in the software package WinBUGS with 15,000 iterations and a burn-in of 5,000. As in previous analyses these burn-in results are discarded. Initial values were set as in section 5.4.3.7 and subjected to sensitivity analysis as in section 5.4.4.2. Subsequent calculation of Incremental net benefit and generation of cost-effectiveness acceptability curves were conducted in the free software programme R. Full WinBUGS and R code are presented in Appendix E

6.5 Results of search for relevant clinical trial literature

The search strategy targeting the link between CDSMP and self-efficacy yielded 2325 abstracts after de-duplication (as a number of databases were searched). After examining these abstracts, 18 papers were ordered. Together with the two studies already included there were 20 studies. Some papers examined a variant of the CDSMP, while others used a non randomised design. Other papers presented too little detail to enable the effectiveness of the intervention to be established. The studies that were included in the analysis are described below. The search strategy targeting the link between self-efficacy and QALYs retrieved no usable data to inform this relationship. Therefore, data from the two sets of IPD were the only evidence available providing evidence of this relationship for the models.

6.5.1 Inclusion and exclusion of studies

This analysis included papers using the CDSMP, or an earlier version the ASMP, as the intervention, which were presented in suitable format (that is a mean effect of treatment over control with a measure of dispersion was either presented or derivable from data presented), and the measure of outcome was not a restricted version of self-efficacy (such as “pain self-efficacy”). A number of “potentially” relevant studies were excluded from the base case for the following reasons:

- 1) the intervention was an arthritis specific version of the CDSMP, the Arthritis Self Management Programme (ASMP),^{74, 77, 213-216} or variant of the CSMP or ASMP,^{73, 78} or an alternative self management programme,¹⁵⁶ and did not use overall self-efficacy as the outcome measure.
- 2) IPD data were available and were already included in the model.^{66, 83}

- 3) Published papers were not identified (only abstracts with minimal data presented).^{217, 218}
- 4) data were not presented in a manner that enabled the effect of treatment on self-efficacy to be established,^{219, 220} for example baseline data were not presented.

6.5.2 Review of included studies

Two studies were available that provided IPD to inform the models described in section 6.4.^{66, 83} These studies have been described in detail in the previous chapter. Additional papers supplied evidence for the models described in section 6.4, and these papers are reviewed below.

6.5.2.1 Papers supplying AD included in analysis

Five papers were identified that provided evidence of the effectiveness of the CDSMP.^{75, 76, 79, 81, 82} All of these papers measured the effectiveness in terms of impact on self-efficacy.

Lorig (2001a), compared the effectiveness of the CDSMP with "usual care".⁷⁹ The intervention was delivered, as is usual, by peer instructors who had been trained in the delivery of the intervention following a "highly structured manual". The study was conducted in a community based setting in the United States (US) with a population of 489 adults (over 18) with one or more chronic conditions. The average (mean) group size was 9.4 with a range from 4 to 18. There was no discussion of the impact of group size on outcomes. The study was not randomised, rather a before and after design was employed. The mean age of the population was 62.2, the majority were married (64%), female (73%) and non-Hispanic white (83%) with an average of 2.3 chronic conditions and 15 years of education.

The results are similar to a randomised trial conducted by the same authors,⁸⁰ though the latter study did not report self-efficacy as an outcome. The former study⁷⁹ showed a "statistically significant" increase in self-efficacy over 12 months, with a mean increase of 0.5 on the 10 point self-efficacy scale (with a standard deviation around the changed score of +/- 2.4).

Clearly, the design of this study means that there is a potential for confounding (a variable that impacts on the outcome that has not been allowed or controlled for). It is unclear whether this intervention is wholly or partly responsible for the improvement in self-efficacy. Nevertheless, given that improvement in self-efficacy is a primary objective of the intervention, this is not unlikely.

Lorig (2001b) again considered the CDSMP delivered by trained peer leaders using a structured manual based in a community setting in the US.⁷⁵ The characteristics of the study population were broadly similar to the study above in that the majority were female (65.4%), married (55.8%), non-Hispanic white (90.8%) with a mean age of 65.3 and an average 2.2 chronic conditions and 14.3 years of education. The average size of the group was not reported. The study design was originally a randomised controlled trial, but results from the study were published as a one and two year longitudinal follow-up (without preserving randomisation). The study reported an increase in self-efficacy of 0.31 (SD 1.67) at one year and this was included in the models below. Again, the study lacked a control group and the results may be questioned due to the potential for confounders.

Three further papers provided evidence of the effectiveness of CDSMP in terms of self-efficacy based on randomised evidence.^{76, 81, 82} Fu et al reported the CDSMP in five urban communities in Shanghai, China.⁸¹ The comparator was a waiting list control and patients were randomised either to immediately receive the intervention or receive it in six months. Age, gender, marital status and number of condition characteristics were similar to the studies above (mean age 64.21, 73.3% female, 82.3% married, 2.09 chronic conditions), but all participants were of Chinese ethnicity. Baseline education levels were slightly different between the groups (9.48 treatment and 9.88 control), but this was adjusted for in the analysis. Baseline levels of self-efficacy was higher in this group (7.36 intervention 7.23 control), and the intervention improved self-efficacy (0.22 over 6 months), while the control groups' self-efficacy level fell (-0.41 over 6 months), providing evidence that the intervention was effective at improving self-efficacy in this patient population.

Siu et al, examined the effectiveness of CDSMP in Hong Kong compared to a Tai-Chi class.⁸² The potential benefits of CDSMP could be expected to be lower in this

study as an active comparator is introduced. As with the studies described above, the study participants were all adults (aged 18 and over) with at least one chronic condition. The design of the study was randomised but in an unorthodox manner. Patients were initially pooled into groups of 20-30 individuals then matched according to age, gender, history of disease and duration of illness and then randomised to CDSMP or Tai-Chi using a coin toss. Presumably this was to ensure approximately the same size groups (which were aimed to be 12-15 for CDSMP and 25-30 for Tai-Chi), though the rationale is not stated. Details of the make-up of treatment and control groups were that participants were mainly female (75%), married (72%) and educated to secondary level (54.1%). The exact make-up of age and history/duration of disease was not stated explicitly, but was reported as “not significantly different” between groups.

Lorig (2003) used a randomised design to explore the effectiveness of the CDSMP in a Hispanic population resident in the US.⁷⁶ The participants were Spanish speakers with a diagnosis of heart disease, lung disease or type 2 diabetes. The control was a waiting list control and both groups were evaluated at 4 month follow-up. As with other studies of this intervention the majority of participants were female (79%) and married (55%). Education level was lower than previous studies by the same author at 7.6 years and mean age and number of conditions were also somewhat lower at 57 years and 1.9 conditions.^{75, 79, 80} Groups were comparable at baseline in all these variables. A majority of both groups (65% intervention and 64% control) were born in Mexico with a small percentage (6% intervention and 5% control) born in the US). Both groups showed improved levels of self-efficacy at 4 month follow-up with the intervention group performing better on this outcome measure.

All these studies showed an improvement in self-efficacy scores in those given CDSMP, though the lack of randomisation in two studies makes the results questionable.

In addition, two papers reported overall self-efficacy from an ASMP intervention,^{71, 73} while one paper reported overall self-efficacy from the CDSMP delivered via the internet.⁷⁸ The ASMP intervention was the forerunner of the CDSMP, and provides

evidence of the link between a lay led self management program and overall self-efficacy. The first paper examined the ASMP in a population of patients suffering from arthritis, again in a sample where the majority were female (79%), and a relatively high level of education (14 years). The study showed the intervention improved self-efficacy, though the before and after design limits confidence in the results. The other study examined the ASMP using a randomised design and again found improvements in self-efficacy, while another study examined the effectiveness of CDSMP delivered via the internet also showed an improvement in self-efficacy. Due to doubts over study design and the transferability of results from ASMP studies, the latter three studies were excluded in a sensitivity analysis to assess their impact on the results.

A summary of the results of these studies is presented in Table 6.2.

6.5.3 Clinical trial summary

Two studies with IPD were available.^{66, 83} The search strategies above identified several trials that provided AD available linking CDSMP with self-efficacy.^{71, 73, 75, 76, 78, 79, 81, 82} No data were available (other than that available from the two studies with IPD available), linking self-efficacy with QALYs or utility. Therefore, in summary we have 4 sources of evidence:

- 1) IPD linking CDSMP and QALY directly*
- 2) IPD linking CDSMP and SE*
- 3) IPD linking SE and QALY*
- 4) AD linking CDSMP and SE*

These sources of evidence can be used in different ways to “model” the relationship between the treatment (CDSMP) and costs (QALYs). These different models structures were described in detail in section 6.3, while the results of the different models are presented below in section 6.6.

Table 6.2 Summary of studies included in the analysis

<i>Study</i>	<i>Date</i>	<i>Intervention</i>	<i>Mean effect size</i>	<i>SD</i>	<i>N</i>	<i>Type of data available</i>
Kennedy ⁶⁶	2006	CDSMP	0.52	0.21	629	IPD
Griffiths ⁸³	2005	CDSMP	0.23	0.13	476	IPD
Lorig ⁷⁹	2001	CDSMP	0.50	2.40	489	AD
Lorig ⁷⁵	2001	CDSMP	0.31	1.67	430	AD
Lorig ⁷¹	1993	ASMP	1.41	2.53	177	AD
Siu ⁸²	2007	CDSMP	0.57	0.39	148	AD
Fu ⁸¹	2003	CDSMP	0.63	0.21	954	AD
Lorig ⁷⁶	2003	CDSMP	0.44	0.30	327	AD
Lorig ⁷³	1999	ASMP	1.14	0.31	286	AD
Lorig ⁷⁸	2006	Internet CDSMP	0.21	0.14	780	AD

6.6 Model Results

Results from the three models are presented below in Table 6.3

Table 6.3 Results from three models with 95% confidence intervals where appropriate

	Direct Model (using IPD only)	Indirect	Mixed
Cost (£)	263 (217 to 306)	193 (188 to 198)	258 (214 to 294)
Qaly	0.009 (0.003 to 0.018)	0.004 (0.001 to 0.007)	0.011 (0.001 to 0.019)
CDSMP effect on costs (β_{CT})	65	N/A	66
CDSMP effect on QALYs (β_{QT})	0.009	N/A	0.007
Impact of self-efficacy on costs (β_{CS})	N/A	-11	-11
Impact of self-efficacy on QALYs (β_{QS})	N/A	0.009	0.009
CDSMP effect on self-efficacy (β_{ST})	N/A	0.474	0.473
ICER	29222	48250	23454

The results are of considerable practical interest. The decision to adopt the intervention may be influenced by the choice of model. The direct estimate using the IPD only estimates an Incremental Cost-effectiveness Ratio (ICER) of £29,222. Current National Institute of Health and Clinical Excellence (NICE) guidance states that interventions delivering a cost per QALY under £20,000 are likely to be an acceptable use of NHS resources, those with an ICER between £20-30,000 per QALY may be cost-effective depending on other considerations, while those with an ICER over £30,000 are unlikely to be cost-effective without strong additional evidence.¹² The direct model, with an ICER of over £29,000 is clearly close to the

accepted maximum, while the indirect model is well above the maximum at almost £50,000 per QALY. The mixed model using all the available evidence produces the lowest ICER at around £23,000. Clearly, the adoption decision based solely on the indirect model would be to reject the intervention as it is unlikely to be cost-effective. Using the direct model with evidence limited to the two UK based RCTs would lead to less certain evidence for the adoption decision, and may be judged cost-effective depending on other factors that may be accounted for. The mixed model has the greatest uncertainty at the threshold values of a QALY currently recommended.

The cost-effectiveness acceptability curves (CEACs), described earlier in Chapters 2 and 4, present a graphical representation around the adoption decision. The CEACs for these models are presented below. These results were similar in the sensitivity analysis where the ASMP studies and the single internet study were excluded. When these studies, which demonstrated an impact of treatment on self-efficacy were excluded, the ICER on the indirect model increased to £48,617. The change is relatively small as the impact of self-efficacy on QALYs is limited. Similarly, excluding the studies with non randomised designs also had limited additional impact, partly due to the aforementioned weak relationship between self-efficacy and the QALY and partly due to two of the three non randomised designs being included in the first sensitivity analysis. The ICER remained at over £48,000 per QALY.

It is interesting to note that the indirect model using the surrogate mechanism via self-efficacy shows the greatest level of certainty at commonly used threshold values of a QALY. This shows that we would be virtually 100% certain (at a cost per QALY of £20,000), that the intervention was *not* cost-effective. Using the other two models, which incorporate the IPD from two UK trials, shows a much greater degree of uncertainty at commonly employed threshold values of a QALY. In fact, the mixed model which incorporates the most evidence is the most equivocal with a 30% chance of the intervention being cost-effective at £20,000 per QALY. At values of £30,000 per QALY, the indirect analysis still shows almost 100% certainty that the intervention is not cost-effective, while the direct model and the mixed model both show that the probability of the intervention being cost-effective exceeds 50%.

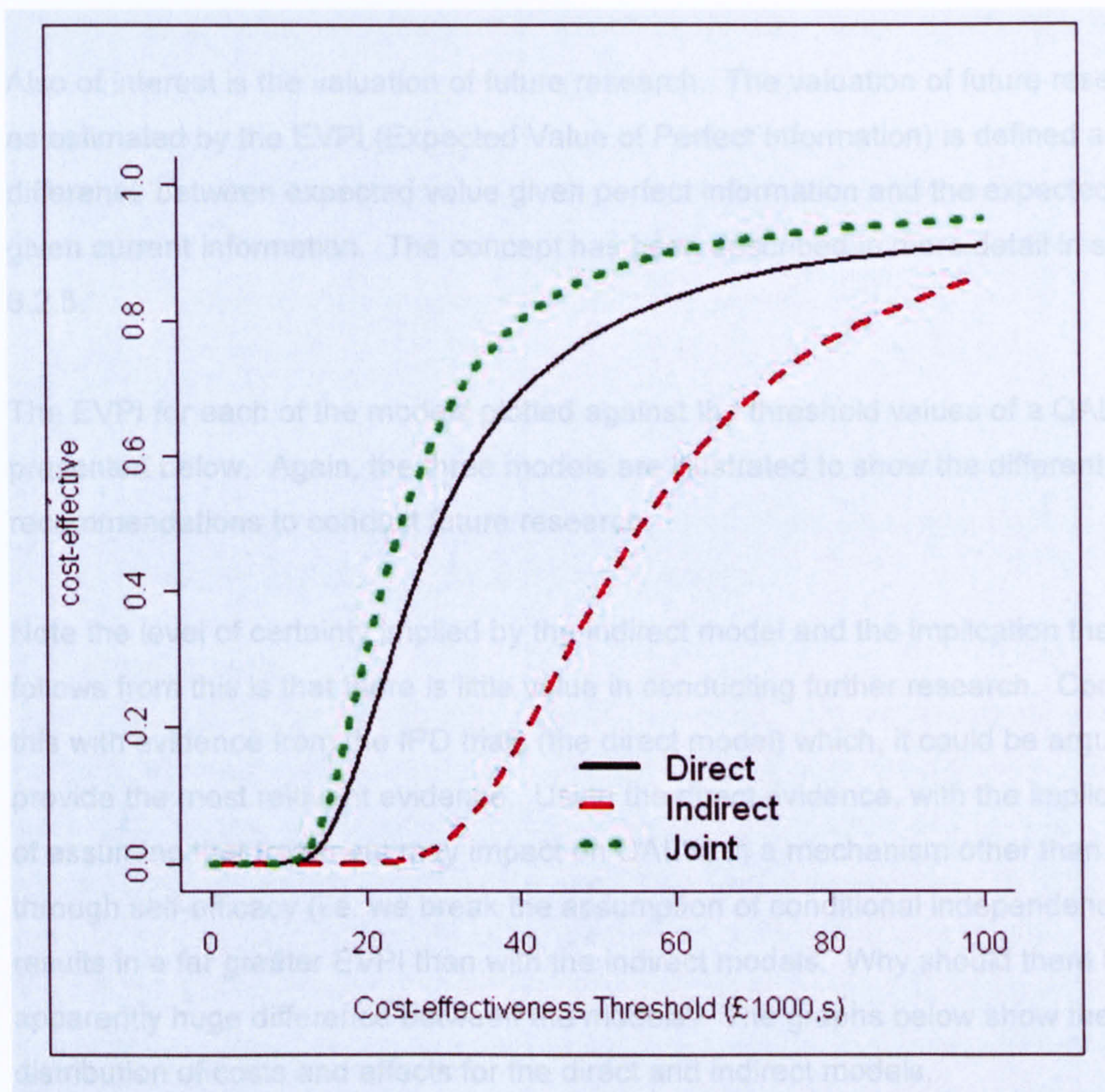


Figure 6.4 CEACs for models using alternative mechanisms

The CEACs demonstrate that there is a clear ordering of the models in terms of the relative cost-effectiveness of the intervention. The indirect model always portrays the intervention to be less cost-effective than the direct model which in turn always portrays the intervention to be less cost-effective than the joint model. Intuitively, we might expect the joint model to lie between the direct and indirect models, but that does not reflect the additive nature of the models. In this instance, the joint model allows information from both direct and indirect sources to impact on costs and QALYs. As there is a positive relationship between treatment and self-efficacy and

self-efficacy and QALYs, in this instance adding the indirect results will lead to the intervention being more cost-effective than simply using the direct results.

Also of interest is the valuation of future research. The valuation of future research as estimated by the EVPI (Expected Value of Perfect Information) is defined as the difference between expected value given perfect information and the expected value given current information. The concept has been described in more detail in section 6.2.5.

The EVPI for each of the models plotted against the threshold values of a QALY are presented below. Again, the three models are illustrated to show the different recommendations to conduct future research.

Note the level of certainty implied by the indirect model and the implication that follows from this is that there is little value in conducting further research. Contrast this with evidence from the IPD trials (the direct model) which, it could be argued, provide the most relevant evidence. Using the direct evidence, with the implication of assuming that treatment may impact on QALYs in a mechanism other than through self-efficacy (i.e. we break the assumption of conditional independence), results in a far greater EVPI than with the indirect models. Why should there be this apparently huge difference between the models? The graphs below show the distribution of costs and effects for the direct and indirect models.

With the direct model, there does not appear to be any strong correlation between costs and effects. Contrast this with the indirect model, where the assumption of conditional independence has induced a very strong correlation between self-efficacy and QALYs, so that if the value of self-efficacy is high, costs and QALYs are likely to be high also. The issue becomes whether this strong correlation is justifiable. It could be argued that as long as this correlation is justified, the indirect model has reduced “noise” (for example it would exclude any unrelated costs and effects that happen to occur in the treatment groups). Alternatively, if the direct effects of treatment on costs, QALYs or both are not equal to zero, we have overstated the certainty.

With the data we had available for these models, we have been able to show that the different models generate different values of the ICERs as well as different probabilities of being cost-effective at various thresholds. More importantly, these models may give differing recommendations for both the adoption decision and the decision whether to conduct further research. This difference in the models is attributable to the differing data used in the models and the assumptions made in the models. In particular, if conditional independence is a valid assumption then the use of the indirect models is possibly appropriate.

However, what constitutes validity in this field is not known. At present, it is a value judgement. The literature in the field of surrogates is not well developed in the area of estimation (as oppose to hypothesis testing) and offers little insight perhaps other than to refute the use of surrogates in these situations.

6.7 Discussion

This chapter aimed to provide a framework for incorporating aggregate data into evidence synthesis and to explore how including these data, identified through a systematic review, might impact on the results and conclusions of analyses. Three models were developed to examine a range of potential mechanisms through which the intervention, CDSMP, might influence final outcomes (in this case, QALYs). The first model, the direct model, examined the effect of treatment (CDSMP) on costs and final outcomes directly, as with a trial based analysis. Secondly, the indirect model examined the impact of treatment on costs and final outcomes indirectly, that is, where all the effects were through the surrogate outcome, self-efficacy. Finally, the third model incorporated both of these mechanisms into the “mixed” or joint model. The alternative models made differing assumptions, but also allowed the use of data from different sources.

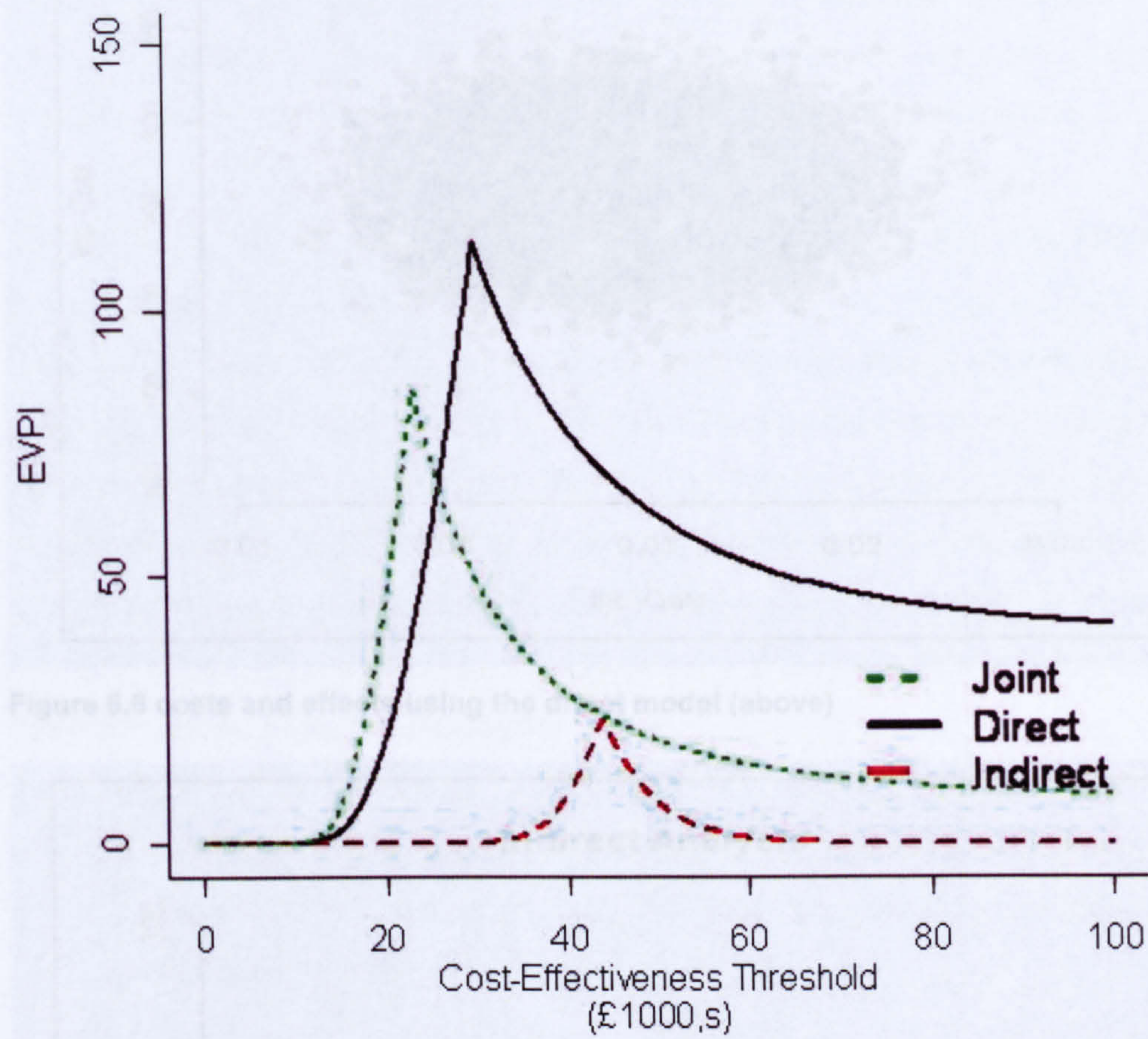


Figure 6.5 EVPI of alternative models

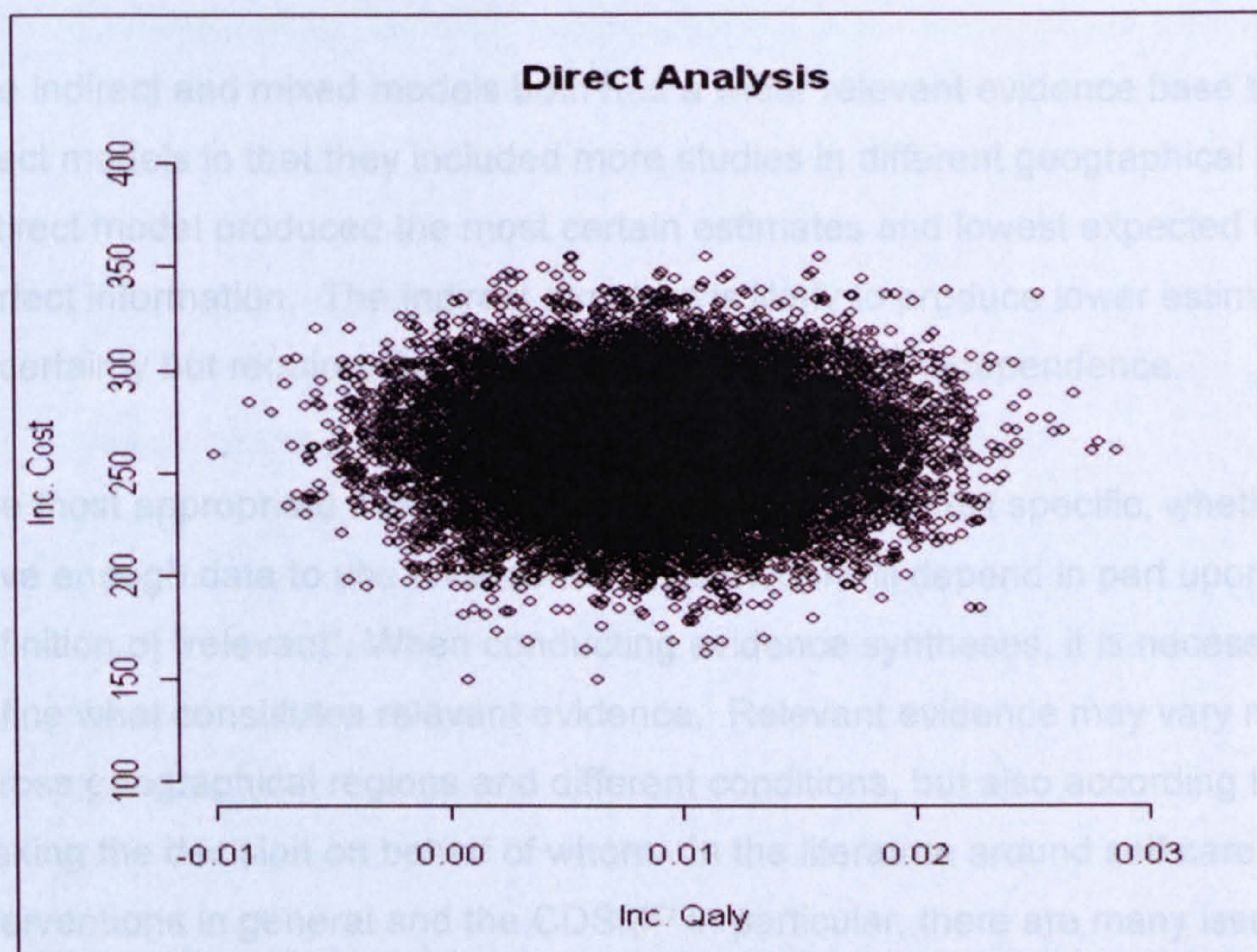


Figure 6.6 costs and effects using the direct model (above)

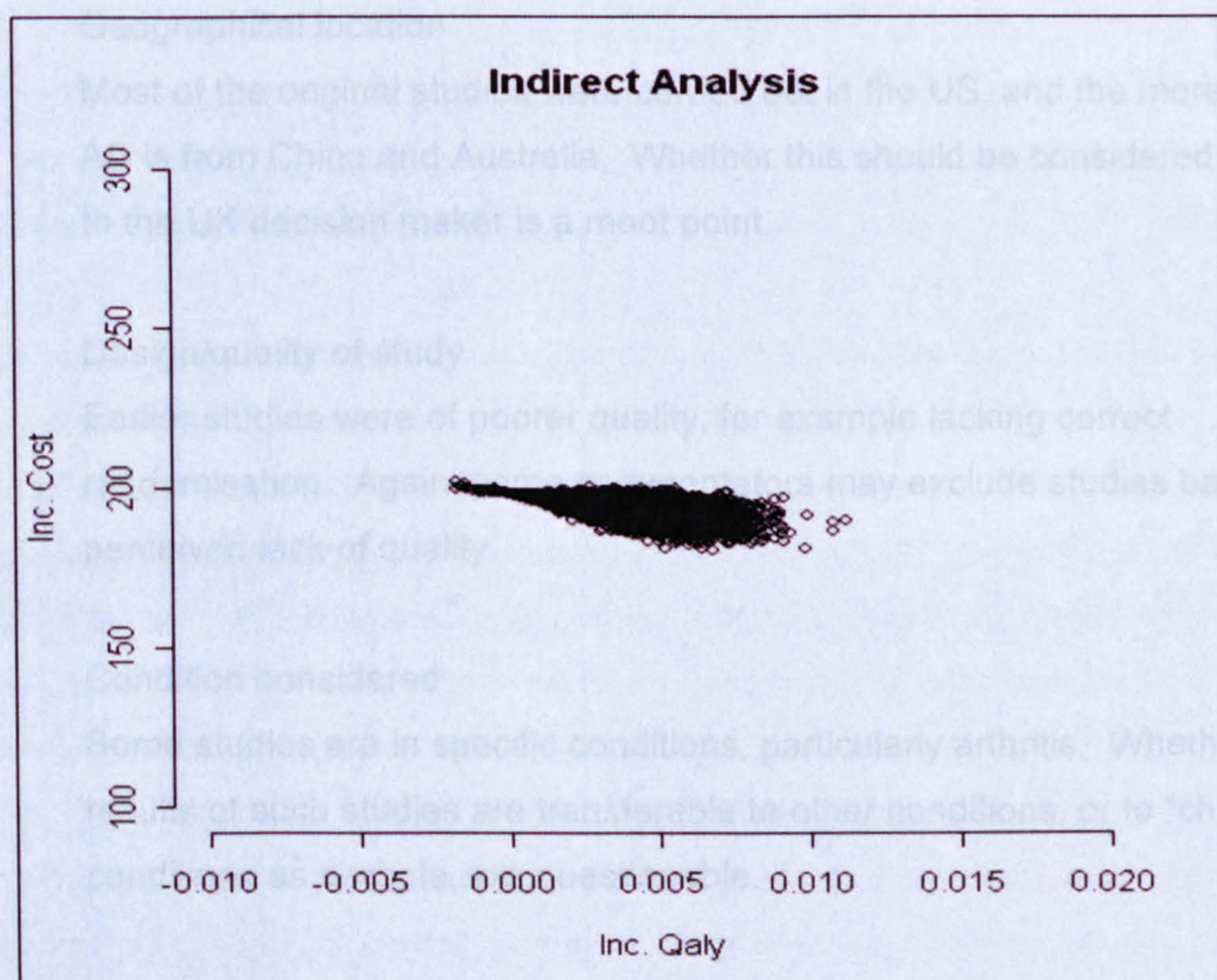


Figure 6.7 costs and effects using the indirect model

The indirect and mixed models both had a wider relevant evidence base than the direct models in that they included more studies in different geographical areas. The indirect model produced the most certain estimates and lowest expected value of perfect information. The indirect structure is likely to produce lower estimates of uncertainty but requires the assumption of conditional independence.

The most appropriate model structure may also be context specific, whether we have enough data to use a random effects model will depend in part upon our definition of “relevant”. When conducting evidence syntheses, it is necessary to define what constitutes relevant evidence. Relevant evidence may vary not only across geographical regions and different conditions, but also according to who is making the decision on behalf of whom. In the literature around self care support interventions in general and the CDSMP in particular, there are many issues around the relevance of types of evidence. These are briefly described below and have been described in an earlier paper.⁷

- **Geographical location**
Most of the original studies were carried out in the US, and the more recent AD is from China and Australia. Whether this should be considered relevant to the UK decision maker is a moot point.
- **Design/quality of study**
Earlier studies were of poorer quality, for example lacking correct randomisation. Again, some commentators may exclude studies based on a perceived lack of quality.
- **Condition considered**
Some studies are in specific conditions, particularly arthritis. Whether the results of such studies are transferable to other conditions, or to “chronic” conditions as a whole, are questionable.
- **Delivery of intervention**

The intervention is usually evaluated in a face-to-face group setting.

However, one study delivered the intervention over the internet. This may or may not constitute a different “intervention”.

- Population evaluated

The CDSMP has been evaluated across a range of ethnic subgroups. The transferability across these subgroups may be limited.

Each of these may be used as reasons to include or exclude evidence as being relevant. For the purposes of this synthesis however, any trial where the CDSMP was evaluated in terms of self-efficacy was considered.

In addition, all studies supplying aggregate data of CDSMP use a hypothesised surrogate outcome measure (self-efficacy) rather than a final outcome measure (such as QALYs) in their analysis. From a decision making viewpoint, it can be argued that all outcome measures other than life expectancy adjusted for quality (as, for example, QALYs) can be considered as “surrogates”, as even “hard” outcomes such as survival must be extrapolated to life expectancy and then either implicitly or explicitly assign a “quality” score to this expectancy. Thus the example of self-efficacy in this analysis can easily be transferred to other economic evaluations with a surrogate outcome.

The relevance and availability of data dictate which model is chosen. In this instance, model choice influences both whether the intervention is likely to be cost-effective and whether it is worth conducting future research. The results are particularly interesting when the different models are compared with each other and/or contrasted with previous chapters. After the single trial based analysis presented in Chapter 4, it may have been reasonable to conclude that this intervention was very likely to be cost-effective at commonly used threshold values of a QALY. However, the evidence presented in Chapter 5 weakened this conclusion, while the recommendations of this chapter are even more equivocal and are entirely dependent on the definitions of relevance of information and the appropriateness of assumptions made. In the absence of the UK trial based analyses, the model using only indirect data would be used. This model would reject the adoption of the intervention and also, crucially, any further research.

Given the evidence we have from the UK based trials, the indirect model understated the uncertainty around the adoption decision and the decision to conduct more research. This is largely due to the assumptions of the indirect model that the intervention worked through self-efficacy. The trial results showed that this model structure was limited and that the intervention influenced costs and QALYs through a mechanism other than self-efficacy.

Though it is accepted that when there are no direct trial based data available, indirect models using surrogates need to be used, care should be taken in their interpretation. If the surrogate is poor, however that is defined, there is likely to be a misrepresentation of uncertainty and potentially misleading evidence on both the adoption decision and the future research decision.

There is no suggestion as to which model is “correct” in the preceding analysis. The correct model is ultimately dependent on our beliefs about the correct structure of the model. The use of model averaging techniques is a potential methodology for examining structural uncertainty. This approach “parameterises” the structural uncertainty and places it directly into the model.⁶⁷ Thus if we have three potential models (as in this case), essentially this approach averages results across the three studies. Results can also be weighted according to their plausibility,⁶⁷ though this requires the valuation of the plausibility of each model. This approach is a promising area for future research.

The analysis provided in this chapter, and the previous two chapters, considered evaluating a specific intervention to support self care using a generic, preference based instrument, the QALY. While this is a commonly used tool, with well understood limitations,^{61, 163} it has been argued that it does not encapsulate all outcomes (health or non-health) that may be of interest to decision makers.²²¹ The following chapter examines the potential for expanding the outcome measure and explores the impact this may have on the cost-effectiveness of the intervention.

6.8 Conclusion

Chapters 4 and 5 demonstrated that the inclusion of additional evidence can impact on results and conclusions. This chapter reinforces this conclusion. Developing

additional models to incorporate additional information requires further assumptions to be made. Using these additional data and assumptions can lead to differing conclusions.

Results and recommendations of the analysis of cost-effectiveness of the CDSMP are sensitive to the type of model chosen and the data that are included. Analysts should be explicit about their choice of model and the assumptions required for that model. Care should be taken when using surrogate or intermediate outcomes as certainty around the decision may be overstated and the value of future research understated as these models require the assumption of conditional independence, which is frequently overlooked.

Chapter 7. What outcomes do patients value? A discrete choice experiment

7.1 Introduction

The previous three chapters have expanded the evidence base in the evaluation of the cost-effectiveness of the EPP from a single UK trial based analysis to an analysis incorporating evidence from a variety of sources and geographical locations. All these analyses considered the outcome of interest to be the QALY. This is a common outcome measure in economic evaluation, though this restricted perspective is not a requirement of economic evaluation in the extra-welfarist paradigm. However, as discussed in Chapter 2, some authors have expressed concern about the use of the QALY as the measure of benefit.^{50, 52} Other authors have argued that the QALY does not capture all outcomes of interest,²²¹ nor does it adequately allow for patient preferences.

Extending the analysis to include other outcomes that are important to patients could potentially impact on the results and conclusion of the analysis. This chapter expands the outcome measure from a simple health related quality of life measure, to a wider measure encompassing other outcomes that patients have stated are important to them and examines the impact that this has on the probability of the EPP intervention being cost-effective. A shorter paper describing the experiment conducted in this chapter is currently in submission with Value in Health.²²² Results from the experiment have been presented at Cochrane colloquia,^{223, 224} and Health Economics Study Group (HESG).²²⁵

7.1.1 Background

Recent policy has targeted self care as a means to improve patient outcomes and reduce costs.^{3, 226} A recent study demonstrated that an intervention designed to enhance patients' ability to self care was effective at improving the self-efficacy of patients, that is, their confidence in their ability to manage their condition.⁶⁶ A separate paper, which also forms the basis of Chapter 4 of this thesis, based on the same clinical trial generated QALY gains for these interventions using the EQ-5D instrument.¹⁶² The EQ-5D instrument measures health related quality of life, across five dimensions namely mobility, ability to self care, ability to perform usual activities, level of pain/discomfort and level of anxiety/depression. QALYs generated in this

analysis did not explicitly include outcomes such as self-efficacy, which was included as an outcome measure in the clinical study, or other important outcomes as described by patients in previous work.

Therefore, while QALYs have a commonly expressed value,¹² they may not incorporate all the outcomes that are of interest. In contrast, while there are other outcomes of interest, we have no knowledge of their “value”. This leads to problems of interpretation as decision makers cannot assess the relative merits of self-efficacy compared to health related quality of life. This chapter aims to address this issue by valuing other “important” outcomes in terms of health related quality of life.

7.1.2 Aims and objectives

One technique that can be used to establish the relative merits of outcomes of interest is the discrete choice experiment (DCE). These experiments are based on individuals stated (rather than revealed) preferences as there is commonly no market for the characteristic of interest to reveal preferences. Individuals are typically asked to respond to a series of hypothetical questions with a number of attributes set at differing levels. For example, a simple question in a simple discrete choice experiment may present an individual with two options A and B. Both have two attributes, for example, access to GP and GPs knowledge of the presenting condition. As an example, option A may be to see the GP immediately but he/she is not very knowledgeable. Option B may be to see the GP in three days time, and the GP is reasonably knowledgeable about the presenting condition. The responder is asked whether he/she prefers A or B. Increasing the number of attributes and levels increases the complexity of the experiment but allows the analyst to gain a fuller idea of the relative merits of the attributes.

This chapter describes a discrete choice experiment conducted to examine the relationship between health related quality of life and other outcomes which may be important to patients with long-term conditions. In addition, the estimation of rates of substitution between the QALY and these other outcomes enable decision makers to include other outcomes in their assessment of cost-effectiveness.

The use of the DCE methodology enables the valuation of important outcomes in terms of QALYs and allows us to “adjust” QALYs to allow for other important measures of outcome (in this example, self-efficacy is chosen). This enables interventions, such as the EPP, to be assessed in terms of both the QALY and also the QALY adjusted for self-efficacy. Ultimately, these additional effects can be incorporated into the cost-effectiveness acceptability curve (CEAC) to demonstrate the impact on the probability of the EPP, in this case, being cost-effective for a range of decision makers threshold value of a QALY. This chapter has a series of aims/objectives which are listed below:

- 1) Can DCEs be used to assess the value of non-health outcomes that can be incorporated into CEA?
- 2) How important are interaction terms in the DCE?
- 3) What problems are encountered using DCEs in CEA and do these impact on the appropriate estimation technique?
- 4) What impact do the results have on the cost-effectiveness of interventions to support self care and what are the policy implications?
- 5) What are implications for DCE in CEA?

7.2 DCEs in the health care evaluation literature

Economic evaluation is conducted because resources are scarce and the demand for the resources will always outstrip the supply. To assess cost-effectiveness, the measure of outcome chosen is some measure of health related quality of life such as the QALY. However, discrete choice experiments have been used as a means of either expanding the measure of outcome or incorporating factors other than health outcomes.²²¹ Discrete choice experiments are a stated preference method used to establish the important constituent parts of a good or service, and are based on the premise (expanded below in section 7.4), that the benefits associated with healthcare interventions can be expressed in terms of:

- a) the *attributes or characteristics* of that intervention,²²⁷ and
- b) the *attributes or characteristics* of the person valuing them.^{228, 229}

For example, speed of access to health care, who provides that health care and where it is provided, are often considered as important aspects of health care that may not be captured by a measure of HRQoL.²²¹ DCEs, or similar techniques, have also been used to generate WTP for QALY,²³⁰ and to generate values for health states for use in economic evaluation.^{231, 232}

DCEs have also been used to estimate preferences in miscarriage management,²³³ management of prostate cancer²³⁴ as well as a variety of other conditions and settings.²³⁵⁻²⁴⁶

DCEs have also been used to estimate willingness-to-pay (WTP) values for a QALY in a Danish population.²³⁰ The estimated WTP of 88,000 Danish Kroner (approximately £8,000) appears low and the authors suggest that this might be due in part to the upper limit presented to respondents being set at too low a level. Doward et al compared the values of health states using DCE and time trade off (TTO) and found that while the ordering of health states was similar between the two methods, actual values were significantly different.²³¹ The authors also found that there was much greater variation within TTO responses and much less understanding of the exercise. Bala and Mauskopf used DCE to estimate the loss function associated with drug treatment for acute myocardial infarction.²³² These authors used welfarism as their theoretical basis and consequently rather than expressing the loss function as a function of QALYs, it is expressed as a monetary amount.

There are few examples in the published literature of DCEs being used to establish marginal rates of substitution between health states and other attributes for inclusion in a multi-attribute utility measure. In this instance, the use of self-efficacy is justified by the importance of this concept in the psychological literature,^{168, 247, 248} where it is an important outcome that would not be wholly encapsulated within a QALY framework.

Similarly, most DCEs in the literature use a limited main effects design (that is, they do not allow for the possibility of significant interaction terms). This analysis considers the importance of the interactions between attributes in the DCE. An example of interaction terms in this discrete choice experiment is presented in section 7.6.2.

The published DCEs in the health economics literature make a number of assumptions about the econometric estimation technique that should be used when analysing these experiments. It is far from clear that previous examples use the correct methodology, and therefore this chapter considers several alternative estimation techniques.

7.3 Design of DCE

Historically, most DCEs in the health care literature have used a limited number of attributes and a simplistic main effects design.²⁴⁹ When relevant attributes, or interaction between attributes, are omitted, biased parameter estimates are likely as respondents may unknowingly assign those values.

7.3.1 Full factorial and fractional factorial designs

A full factorial design represents the full range of combinations of attributes and their levels. For example if we had 4 attributes with 3 levels each (as in the DCE described below), this would require 3^4 or 81 scenarios (or profiles) would be generated. Such designs are practical only for experiments with limited numbers of attributes and attribute levels.²⁵⁰ Full factorial designs for large numbers of attributes/levels would present considerable cognitive burden on respondents. As with previous DCEs in health and other disciplines, a fractional factorial design was chosen. Fractional designs select a subset of the full factorial, but allow the relatively efficient estimation of the coefficients around attributes.²⁵⁰ These designs inevitably lose some information when compared with the full factorial design, but reduce the cognitive burden for respondents.

7.3.2 Properties of DCE design

Orthogonality, level balance, utility balance and minimal overlap are considered to be appropriate characteristics of DCEs.²⁴⁹ Orthogonality describes the property that each level of an attribute should appear the same (approximately) number of times with each other level of each other attribute. For example, health state level 1 should not appear solely in the same choice set with confidence level one, but should appear equally distributed with confidence levels 1, 2 and 3. Level balance implies that the levels each attribute should appear with equal frequency, while minimal overlap implies that the instances where the attribute level is the same in a given choice set should be minimised (thus for example, if option A has health state 1, option B should ideally not have health state at level 1). Utility balance implies that each choice set option should have similar utility and therefore similar probability of being chosen. A trade-off exists with this characteristic as clearly making choices more difficult may increase the cognitive burden and reduce response rate.

To ensure the most efficient statistical design, the number of levels of each attribute should be multiples of each other. That is, if we have 4 attributes, we could have 3 with 4 levels and 1 with 2 levels, or even 1 with 8 levels. We should not, however, have some attributes with 2 levels and others with 3. The practical implication for this DCE is that there should be 4 attributes each with 3 levels (as having three levels for each clearly satisfies the above requirement). Not only is this statistically most efficient, but it also enables some 2 way interactions to be included in a 27 question design (see 7.3.3 below).

From the orthogonal design described above, choice sets (such as the one presented in Figure 7.1) can be created using the “foldover” technique. Consider a 4 attribute 3 level experiment such as the one described in this chapter. Using a catalogue design,^{251, 252} numerous scenarios can be developed. The first scenario generated may have *attribute 1* at *level 1*, *attribute 2* at *level 2*, *attribute 3* at *level 2* and *attribute 4* at *level 3*. We can represent this as 1223. This profile would be one scenario in the choice set. To generate the comparison, which responders will

choose between, levels are “folded over”. Thus, for example, level 1 becomes level 2, 2 becomes 3 and 3 becomes 1. Therefore, 1223 becomes 2331 when folded over. The comparator then becomes *attribute 1 at level 2, attribute 2 at level 3, attribute 3 at level 3, and attribute 4 at level 1* (or 2331). This creates one choice set. The procedure is repeated for the number of questions to be asked to create the full questionnaire.

7.3.3 Interaction terms and length of questionnaire

In recent years the DCE literature has seen more emphasis on the design of experiments. Issues addressed include the statistical efficiency of the designs,²⁴⁹ responder efficiency,²⁴⁹ completeness of preferences and the use of simplifying heuristics,^{253, 254} and the meaning of willingness-to-pay (WTP) measures in DCEs.²⁵⁵⁻²⁵⁸ The structure of the choice model has received less attention. Most experiments have used a linear additive model due to the additional complexity and cognitive burden for responders of including many attributes and allowing for interaction terms. This is justified by the frequently quoted result that between 70 and 90% of explained variance can be accounted for by main effects models,²⁵⁹ that is, models that do not allow for interactions between attributes. However, these results are dated and should not preclude the inclusion of two way interaction terms, particularly where there is an *a priori* expectation that interactions may be important. Kocur et al list three consequences of assuming interaction terms are negligible namely: these interactions remain undetected, experimental error is increased and incorrect conclusions may be drawn about significance of certain attributes.²⁵¹ Louviere et al suggest the use of an endpoint design to allow for interaction terms;²⁵⁰ unfortunately, the use of this technique requires that the levels within the attributes are on an interval scale i.e. that the movement from health state 1 to health state 2 is identical from the move from 2 to 3; with the health state dimension in this study, this is unlikely to be the case. Movements from 2 to 3 are associated with much greater utility decrements than movements from 1 to 2.

Some authors have examined the use of interactions (between attributes) within a DCE.²⁶⁰ Typically allowing for interaction terms increases the number of scenarios the responder has to answer several times. For example, a DCE with only 3

attributes each with 3 levels would require only 9 questions if interaction terms are ignored. However, to include all the interaction terms requires 27 questions. As the number of attributes and levels increases the complexity of a DCE including interactions increases. However, there is no consensus as to the “correct” number of questions to ask individuals, with previous work in the DCE field suggesting that length of questionnaire has little impact on response rate.²⁶⁰

While there is little evidence of the importance of interactions in this field, it was felt that interactions between particular attributes could be important and therefore, this DCE used a design enabling the estimation of interaction terms for a subset of the sample.

7.4 Theoretical basis for DCEs

The standard economic model of discrete choice experiments is based on *Random Utility Theory* (RUT) and the *Lancaster's characteristics theory of value*, whereby the indirect utility of consuming a good is a function of the characteristics or attributes of that good and also the individual's characteristics. That is, the utility of good A to individual i is dependent only on the characteristics of good A (x_A) and the characteristics of the individual (s_i), or more formally $U_{Ai} = U(x_A; s_i)$. This function is then decomposed into a potentially observable deterministic component of utility V_{Ai} and a random element ε_{Ai} such that $U_{Ai} = V_{Ai} + \varepsilon_{Ai}$.

For simplicity $V(.)$ is often assumed to have a simple additive linear functional form, such that $V_{Aj} = \beta_0 + \beta_1 x_{1A} + \dots + \beta_N x_{NA} + \gamma_0 + \gamma_1 s_{1i} + \dots + \gamma_M s_{Mi}$. Where x_{nA} is the value of the n th characteristic of good A and s_{mi} is the value of the m th characteristic of individual i. β_n and γ_m are the individual invariant (unobserved) parameters related to x_{nA} and s_{mi} , which transform the characteristics of the services and the individual into a utility of that service to that individual. The objective of the DCE is to obtain estimates for the vector of unknown parameters, $(\beta_0, \beta_1, \dots, \beta_N)$ (the part-worth utilities). The ratio of any two of these parameters gives the rate of substitution between the variables to which they are associated. Thus if say β_j is the parameter

associated with a cost dimension (money, time or distance) and β_1 is the parameter associated with a higher health state, then β_1/β_5 represents the relative rate of substitution between the two and therefore may be used to value variables in terms of a cost. ε_{Ai} is assumed to be the standard unobservable error term with the usual properties i.e a random draw from a defined distribution with a mean of 0 and a constant but unknown variance (i.e. homogenous). The error terms are also assumed to be independent of the observed characteristics.

In a DCE where I individuals are asked to make J discrete choices between paired profiles A and B , the paired profiles consists of two goods or services completely described by the vector of variables x_{jA} and x_{jB} . The levels and values of the characteristics of the profiles are varied in each choice according to some experimental design in order to create sufficient variation. Any possible omitted variable or characteristic of each profile is explicitly stated to be constant across all profiles, such that all differences between profile A and profile B in choice j are observed and captured by differences between the vector of variables x_{jA} and x_{jB} .

In each choice the rational respondent will choose the option that yields the highest level of utility to them and so in choice j individual i will choose B over A if $U_{jBi} > U_{jAi}$. From the construction of these indirect utility functions it can be shown that the difference in utility is due to the difference in attributes between option A and B in choice j and the error term.

$$\begin{aligned}
 y^*_{ij} &= U_{jBi} - U_{jAi} \\
 &= (\beta_0 + \beta_1 x_{jB} + \dots + \beta_N x_{jN} + \gamma_0 + \gamma_1 S_{1i} + \dots + \gamma_M S_{Mi} + \varepsilon_{jBi}) - \\
 &\quad (\beta_0 + \beta_1 x_{jA} + \dots + \beta_N x_{jN} + \gamma_0 + \gamma_1 S_{1i} + \dots + \gamma_M S_{Mi} + \varepsilon_{jAi}) \\
 &= \beta_1 \Delta x_1 + \dots + \beta_N \Delta x_N + \varepsilon_{jBi} - \varepsilon_{jAi} \\
 &= \beta_1 x_{1j} + \dots + \beta_N x_{Nj} + \varepsilon_{ji}
 \end{aligned}$$

where $x_{1j} = x_{1B} - x_{1A}$ and $\varepsilon_{ji} = (\varepsilon_{jBi} - \varepsilon_{jAi})$

As specification individual characteristics appear in equal measure on both sides of the difference equation, they simply difference out and hence the decision to choose A over B or B over A is independent of individuals' observable characteristics.

7.5 Econometric estimation techniques for DCE

The independent variables were the differences in the levels of the variables. The independent variables were dummy coded, with the “best” level of each attribute the omitted category. “Best” in this instance refers to *a priori* expectations that higher levels of health status and confidence would be preferred to lower levels, more frequent visits from friends/relatives would be preferred to fewer and that quicker GP access would be preferred to slower. Table 7.1 shows each attribute and its levels. The omitted level in each case was level one (only health problem is moderate pain/discomfort, totally confident, can see GP tomorrow and see friends/relatives daily). The results described below and in Table 7.4, show decrements in utility associated with movements from level one of the attribute to levels two and three.

The dependent variable y_{ij}^* is unbounded and continuous, it may be positive indicating a preference for B over A, or negative indicating a preference for A over B. It may also be zero which would indicate indifference between the two choices. However the analyst does not observe this latent variable but instead observes the outcome, y_{ij} - a binary outcome equal to 1 if option B is chosen and 0 if A is chosen.

Thus;

$$y_{ij} = \begin{cases} 1 & \text{if } y_{ij}^* > 0 \\ 0 & \text{otherwise} \end{cases}$$

Taking the previous equation and substituting in for y_{ij}^* , then $y_{ij} = 1$ if $\sum \beta X + \varepsilon > 0$

i.e. $y_{ij} = 1$ if $\sum \beta X > -\varepsilon$. And so the probability that option B is chosen is given by:

$$prob(y_{ij} = 1) = prob(\sum \beta X > -\varepsilon).$$

In order to estimate such a model some assumption is required about the distribution of ε and different assumptions about this distribution define the logit and

probit models. For example if we assume the error term is iid $\sim N(0,1)$, (independently and identically distributed with a mean of zero and a constant variance) a draw from the standardised normal distribution, then the probit model is assumed. Given an assumption on the behaviour of the error terms, a set of values X and a set of 1 or 0 outcomes, the β 's may be estimated. For example, suppose we had 1000 observations, split evenly 500 ones and 500 zeros, then the estimated constant term would be $\hat{\beta}_0 = 0$.

Returning then to the context of the DCE, suppose we give an individual a choice between two identical options A and B, then we would theoretically expect the individual to be indifferent between them i.e. a 50/50 chance of choosing B over A or vice versa. Thus we would clearly expect $\hat{\beta}_0 = 0$ (or at least not significantly different from zero). Any value other than zero, indicates a systematic preference for A over B ($\hat{\beta}_0 < 0$) or B over A ($\hat{\beta}_0 > 0$) all other things being equal. In practice, since on a paper questionnaire, B appears on the right hand side and A on the left, it is plausible that a predominantly right handed population may well systematically choose B over A and give a positive non-zero constant term. However, remember that the β 's may be compared to each other to establish relative values and trade-offs between dimensions. Thus a relatively large constant term implies respondents are willing to trade substantive benefit for not having to exert the extra effort to tick a left sided box. In other words you would expect there to be a lexicographic preference for all attributes over the constant term. That is, that improvement in any of the attributes should not be outweighed by the constant term. Evidence to the contrary suggests there is something suspect about the underlying premise of the exercise.

The role of the constant term raises its head again when dealing with the fact we observe many observations/choices from individuals. This causes issues with the assumption of independent error terms and the standard response to this in health economics appears to be implementation of a random effects model. However, random effects models account for correlation by allocating each individual their own choice-invariant constant term. Thus random effect models simply allow for

heterogeneous preferences for choosing B over A all other things being equal. Thus it is a rather curious solution to a problem as it is inconsistent with the underlying theory and hence it is arguably more of a specification test than a genuine solution.

To summarize the above, it is argued that: an empirically significant constant term is inconsistent with the underlying theory and is a cause for concern and could be evidence of model misspecification and thus, far from being a solution to correlated error terms, significant random effects simply imply a heterogeneous rejection of the underlying model.

We illustrate these arguments with our empirical models following the methods section.

7.6 Methods

DCE, a questionnaire based stated choice method, was used to explore the attributes (or characteristics) that are most valued by patients. There are several stages to conducting a DCE, and these stages are described below.²⁶¹

7.6.1 Defining attributes and their levels

Qualitative interviews and focus groups carried out alongside a randomised controlled trial of the Expert Patients Programme (EPP) and reviews of the published literature were used to identify the attributes that patients valued. Based on these sources, three outcomes were selected for inclusion into the study; access to General Practitioners, level of social isolation and level of self-efficacy (patients' confidence in their ability to manage their condition). In addition, because the rationale of the study was to assess the importance of these other outcomes relative to HRQoL and examine whether such outcomes could be included in cost-effectiveness analysis, a measure of HRQoL was included. As it is frequently used in economic evaluation and has previously been used in DCEs,^{230, 249} the EQ-5D was used as the basis for measuring HRQoL. EQ-5D measures patient health status across five dimensions (mobility, self care, usual activities, pain/discomfort and anxiety/depression) with three possible responses (no problems, moderate problems or severe problems) for each dimension. This locates each participant

into one of 245 mutually exclusive health states (with the additional states being death and unconscious). Clearly, including this number of levels of an attribute is impractical. Three states were selected as levels for this attribute to maintain the statistical efficiency of the experiment. The three states showed a clear ranking between them and were all levels that were considered plausible for this patient group.²⁶² The attributes and levels identified from this process are presented in Table 7.1.

Table 7.1 Attributes and levels used in the DCE

Attribute	Levels
Health Related Quality of Life	1 No problems with mobility, usual activities, self care or anxiety/depression. Moderate pain/discomfort 2 No problems with usual activities, self care or anxiety/depression. Some problems with mobility and moderate pain/discomfort 3 No problems with usual activities. Some problems with mobility and self care. Also moderate pain/discomfort and moderate anxiety/depression
Level of confidence	1 Totally confident in ability to manage condition 2 Moderately confident in ability to manage condition 3 Not at all confident in ability to manage condition
Access to General Practitioner	1 GP appointment tomorrow 2 GP appointment in one week 3 GP appointment in 3 weeks
Level of isolation	1 See friends/relatives daily 2 See friends/relatives every few days 3 See friends/relatives rarely

As described in section 7.3.2, the number of levels in each attribute should be multiples of each other.²⁵⁰ Having more than three levels for each attribute increases the number of questions responders are asked and therefore increases the cognitive burden considerably.

7.6.2 Questionnaire methods

Two questionnaires were developed to test for both differential response rates and to allow for interactions between attributes. The longer questionnaire contained 28

questions and allowed for interactions on 3 of the 4 attributes, but posed a considerably larger cognitive burden on responders. The shorter questionnaire consisted of 10 questions, but only considered main effects (i.e. it did not allow for estimation of interactions between attributes). Both questionnaires are presented in Appendix F.

Therefore, the use of differential lengths of questionnaires allowed us to test for differences in response rate but primarily to assess the impact of interaction effects between attributes. For example, if the interaction effects are ignored, bad health and low confidence each account for a reduction in utility, but it is possible that these two combined reduces utility by more than the sum of the two. For example, being in a bad health state may reduce an individuals' utility by 0.2 and having low levels of confidence may reduce utility reduces by 0.1. However, having both bad health and low confidence may result in a drop of utility greater than 0.3. Most DCEs ignore interactions as they add to complexity and to the cognitive burden for respondents.

Hypothetical choice sets (i.e. a comparison of alternatives) were created using a fractional factorial design with foldover. As described above in section 7.3.2, the first scenario was generated from a catalogue of designs²⁵² and had health state at level 1, self-efficacy at level 2, GP access at level 2 and isolation at level 3. This can be represented as 1223 and would form one scenario in the choice set. Foldover, as described in section 7.3.2, was then used to generate the comparator. The original scenario (1223) and the comparator (2331) form one choice set which appears as one question in the questionnaire. This procedure was repeated to generate the full questionnaire. This design is commonly used in DCEs in health and other disciplines, as it allows for a relatively efficient estimation of the coefficients around attributes.²⁵⁰ These designs inevitably lose some information when compared with the full factorial design²⁵⁰ (where the full range of combinations of attributes and their levels are presented), but reduces the cognitive burden for respondents.

Each choice set required the respondent to select one of the unlabelled options A or B (i.e. it was a "forced choice"). Some authors,²⁶³ have suggested that additional

options should be included (such as an opt out clause) so that the experiment is not a forced choice, while others,²⁶⁴ have claimed that this may increase the number of neutral responses and thereby the “power” of the study. Pilot work on this study indicated that neutral responses were likely in this DCE and a forced choice was chosen as appropriate.

The design of both questionnaires was orthogonal, but no attempt was made at utility balance as this increases the complexity for responders (and also requires *a priori* knowledge of individual preferences). An example of the questions facing patients is presented in Figure 7.1.

7.6.3 Consistency

One question in each questionnaire (short and long) was replicated to check for consistency in responses. Where the responder gave different answers to the same question, this was considered “irrational”, for the purposes of this analysis. There is a large literature on testing for rationality of responses in DCEs (see for example^{255, 265-269}), and a degree of controversy over what constitutes “irrationality”, how to test for rationality and how to proceed with “irrational” responses.

As there is a lack of consensus on testing for, and the appropriate method for dealing with, “irrational” responses,^{265, 268, 269} analyses including and excluding these “irrational” responses were conducted.

When this additional check for consistency was added to the questionnaires, the short questionnaire contained 10 questions (9 different), while the long questionnaire contained 28 questions (27 different).

Imagine that you can have either option A or option B, which would you choose? If you would choose the option where you have **moderate** pain or discomfort, are **not** confident that you can manage your condition but you can have a GP appointment **tomorrow** and you see friends or relatives **daily** (i.e. everything in column A) then choose option A.

However, if you would prefer the option where you have **moderate** pain or discomfort as well as having **some** problems with walking, but you are **totally** confident that you can manage your condition with a GP appointment in **3 days** time, but you **rarely** see friends or relatives (i.e. everything in column B), then choose option B.

<p><i>A</i></p> <ul style="list-style-type: none"><input type="checkbox"/> You have no problems walking about no problems with self care no problems with usual activities moderate pain or discomfort no anxiety or depression <input type="checkbox"/> You are not confident you can manage your condition<input type="checkbox"/> You can have a GP appointment tomorrow <input type="checkbox"/> You see your friends or relatives daily	OR	<p><i>B</i></p> <ul style="list-style-type: none"><input type="checkbox"/> You have some problems walking about no problems with self care no problems with usual activities moderate pain or discomfort no anxiety or depression <input type="checkbox"/> You are totally confident you can manage your condition<input type="checkbox"/> You can have a GP appointment in 3 days' time <input type="checkbox"/> You rarely see friends or relatives
--	-----------	--

Please tick one box:

Choice A

☐

Choice B

☐

Figure 7.1 Example question from discrete choice experiment

7.6.4 Pilot study

Questionnaires were tested via a pilot study. Issues of interest included whether patients understood the questionnaire, whether the attributes were traded, whether one attribute dominated the others, and whether the responses were internally consistent. For the pilot study, 27 individuals with a chronic health problem completed a questionnaire and were then telephone interviewed. In general, patients understood the exercise, though the hypothetical nature of the experiment was not emphasised sufficiently. The hypothetical nature of the study was stressed more for the main study questionnaire. Several responders felt that the long questionnaire was repetitive, but most patients agreed with the choice of attributes. Some responders identified other attributes which they felt may be important such as continuity of care and the quality of the interaction with health professional(s). Several responders stated that they would have opted out of the choice where there was little perceived difference between the options offered, if they had been given an opt out option. These respondents were ultimately able to make a choice between the options, and for this reason the opt-out was excluded from this questionnaire. The disadvantage to the omission of the opt-out is that there is a possibility of overstating the importance or relative weight of attributes in the DCE.

7.6.5 Main study

Postal questionnaires were posted to a sample of 511 chronically ill people (255 short questionnaires, 256 long questionnaires). Patients were randomly allocated to receive either a short or a long questionnaire and a freephone number was provided for any questions patients might have had. Patients who did not respond after two weeks received a written reminder.

7.6.6 Main study sample

Patients who were involved in the randomised controlled trial of the Expert Patients Programme described in Chapter 4 and elsewhere,^{66, 162} and had not indicated that they would prefer not to receive any more questionnaires, were included in the study. Thus the study sample was patients with a (self defined) chronic condition. There were no exclusion criteria and patients from all 28 Strategic Health Authorities were included.

7.6.7 Model estimation

The DCE was analysed by treating each choice between pairs as a single observation. Participants therefore provided either 10 observations if they completed the short questionnaire or 28 observations if they completed the longer questionnaire. The participants' response to each question (i.e. A or B) was included in the model as the binary dependent variable. The independent variables were the differences in the levels of the variables. "Correct" standard errors, allowing for correlation between observations from the same individual were included by clustering on patient identification number.

In the first instance, a standard probit model with constant was used. Subsequently, a random effects probit model with constant was employed with clustering to allow for multiple responses from the same participant. The inclusion of constant terms is a violation of the theoretical basis of the model (see above), but can be used as a notional misspecification test.

The continued identification of a significant constant term lead to other models being considered. There is little guidance in the health economics literature on the procedures for dealing with a significant constant; the problem seems to be largely ignored.

7.6.8 Incorporating the estimates of the value of self-efficacy into the CEAC

The methodology described above can be used to generate values for any of the outcomes in the DCE in terms of EQ-5D scores. Health states 1, 2 and 3 are associated with utility scores of 0.796, 0.727 and 0.552 respectively, from the EQ-5D tariff.¹⁷⁴ Thus the difference between one year spent in health state 1 compared to one year in health state 3 is 0.244 of a QALY. Using the example of self-efficacy, we can see that the coefficient around the movements between health states and self-efficacy levels can be compared (i.e. a marginal rate of substitution (MRS) can be calculated) and applied to this figure to estimate the additional impact of self-efficacy. Thus if the coefficient around the movement from self-efficacy level 1 to self-efficacy level 3 were 0.2 and that around the movement from health state 1 to health state 3 were 0.4, then crudely, health state movement is valued twice as

highly as self-efficacy ($MRS=2$), and the value of spending a year in self-efficacy level 1 over self-efficacy level 3 can be estimated as 0.122 (that is, $0.244/2$).

Using the data from the EPP trial, patients were categorised into low, moderate or high self-efficacy. Scores of one, two and three on the self-efficacy outcome measure were considered “low”, while four, five and six were considered moderate, with scores of seven to ten considered “very confident”. Thus movements from “low” to “very” confident (and all other combinations), could be categorised. These movements between self-efficacy categories can be valued in terms of QALYs (as above). Then, assuming that self-efficacy and QALYs are completely independent, the effectiveness of interventions can be adjusted for the impact of self-efficacy by simply adding this additional amount. While this is a crude adjustment, this estimate reflects the maximum impact that self-efficacy could have, as it is likely that there will be some overlap between QALYs and self-efficacy, this analysis assumes that this overlap is zero.

7.6.9 Sensitivity analyses

A number of sensitivity analyses were conducted. While the majority of the health economics literature has excluded “irrational” responses, this has been questioned by some authors^{265, 268, 269} who suggest that there may be a number of alternative explanations for these responses. Therefore, a standard random effects probit was employed to test whether the inclusion of these responses altered the results or conclusions.

Two questionnaires were administered. It is possible that responses to the two questionnaires were systematically different. Thus a standard random effects probit was used on both questionnaires to identify any potential differences.

7.7 Results

511 patients who had participated in the RCT of the EPP were sent a postal questionnaire. Of these 367 (71.8%) completed and returned one of the questionnaires. Responders were, on average, slightly younger, more likely to own their own home and be in paid employment compared to non-responders. The

characteristics of patients who responded compared with non-responders to the questionnaire are presented in Table 7.2. Response rate among those who were sent the shorter questionnaire was higher than for the longer questionnaire (73.7% vs 69.9%), though not substantially so. The characteristics of those returning both types of questionnaire are also presented in Table 7.2.

Patients’ whose responses were inconsistent (in that they gave different answers to the repeated question) were excluded from the primary analysis (n=98, 26.7%). Surprisingly, there were more inconsistent responders to the short questionnaire (n=64, 34.0%), than the long questionnaires (n=34, 19.2%). The remaining 269 responses were considered in the primary analysis.

Variables are dummy coded, with the “best” level of each attribute the omitted category. “Best” in this instance refers to *a priori* expectations that higher levels of health status and confidence would be preferred to lower levels, more frequent visits from friends/relatives would be preferred to fewer and that quicker GP access would be preferred to slower.

Table 7.2. Characteristics of non-responders and responders type.

	<i>Non-responders (n=144)</i>	<i>Responders (n=367) (average)</i>	<i>Responders (n=188) (short quest)</i>	<i>Responders (n=179) (long quest)</i>
Characteristic				
Age	52.5	57.5	56.9	58.0
% female	72.9	68.7	73.9	63.1
<i>Accommodation status:</i>				
Owner occupied	57.6	74.4	72.9	76.0
Rented from LA	31.9	18.0	18.1	17.9
Rented privately	8.3	6.0	6.4	5.6
Other	2.1	1.6	2.7	0.6
<i>Condition:</i>				
Musculoskeletal	31.9	37.6	39.9	35.2
Endocrine	13.2	11.2	10.1	12.3
Circulatory	8.3	6.5	4.8	8.4
ME	6.3	7.4	6.9	7.8
Other	40.3	37.3	38.3	36.3
<i>Employment status:</i>				
Employed	16.0	20.4	18.1	22.9
Retired	33.3	39.8	36.7	43.0
Unable to work	36.8	31.6	37.8	25.1
Unemployed	6.3	3.5	4.3	2.8
Other	7.6	4.6	3.2	6.1

7.7.1 Interaction terms

The design of the long questionnaire permitted the estimation of interaction terms between attributes. The results of this analysis are presented in Table 7.3. Several interaction terms were dropped due to collinearity. Few of the remaining interaction terms had a substantial impact on the results. Three of the “significant” results were interactions between GP access and health. It is not clear *a priori* which direction these should take, though it is perhaps surprising that these are all negative, thereby implying that speedy GP access even in relatively poor health states is not very important. Noticeably, the interaction between health level 3 and self-efficacy is positive, indicating that if patients are in a relatively poor health state, improving their level of confidence improves their level of utility.

Table 7.3. Interaction terms in the DCE

Attribute/interaction	Coefficient	Standard Error
Health 2	-0.084	0.042
Health 3	-0.496	0.074
Self-efficacy 2	-0.081	0.029
Self-efficacy 3	-0.452	0.048
Isolation 2	-0.016	0.025
Isolation 3	-0.497	0.051
GP access 2	-0.080	0.022
GP access 3	-0.267	0.039
<i>Interactions</i>		
Health 1/self-efficacy 1	-0.041	0.035
Health 1/self-efficacy 2	0.034	0.028
Health 2/self-efficacy 3	-0.033	0.026
Health 3/self-efficacy 1*	0.050	0.029
Health 1/GP access 1**	-0.065	0.025
Health 2/GP access 1**	-0.099	0.034
Health 2/GP access 2**	-0.080	0.030
Health 3/GP access 3**	0.084	0.022
Self-efficacy 1/GP access 3	-0.005	0.030
Self-efficacy 2/GP access 2	0.000	0.022
Self-efficacy 3/GP access 1	0.017	0.026
Self-efficacy 3/GP access 3	0.015	0.033

* significant at 10%, ** significant at 5%

7.7.2 Standard probit model with constant term

Results from this analysis are presented in Table 7.4. The coefficients of each attribute are based on the short and long data combined and represent the impact of a unit increase of each attribute on the probability of getting one outcome (with all the other variables at their mean), and thus the relative importance (marginal rates

of substitution) of each attribute can be estimated by dividing one coefficient by another.

The coefficients reflect the disutility associated with moving from one state to another and are intuitively appealing in that the coefficients move in the anticipated direction. The only exception is the movement from isolation level one (“friends and relatives visit daily”) to isolation level 2 (“friends and relatives visit every few days”), where the latter is preferred. This movement is not statistically significant and is also plausible in that many individuals with chronic illness would find daily visits burdensome.

As expected, the movement from level one to level three is also greater than movement from level one to two for all attributes.

The coefficients around health status, self-efficacy (confidence) and, to a lesser extent isolation, are of similar magnitude. In particular, the results indicate that patients value a movement from health state one (moderate pain, but no problems on other dimensions) to health state three (moderate pain, moderate anxiety/depression, some problems with self care and some problems with mobility, no problems with usual activities) as approximately the same as a movement from confidence level one (“totally confident in ability to manage condition”) to confidence level three (“not at all confident in ability to manage condition”) (as the ratio of confidence level 3 to health state 3 is 1.013).

A difference in “utility” between health states one and three of 0.244 can be generated from the EQ-5D tariff⁸. Thus the movement from not at all confident to totally confident would equate to a gain of 0.25 QALYs if maintained for one year (and assuming that we can ascribe utility values from patient generated responses rather than those of the general public).

Separate analyses for short and long questionnaires were consistent with the combined analysis.

Notice that the constant term included in this model is significant. This constant has no substantive meaning in an unlabelled experiment. In essence, this result implies that responders prefer A to B, even accounting for differences in the scenarios presented to them. Indeed, the size of the coefficient around the constant leads to a more worrying conclusion. If scenarios A and B are identical except that B has health state at level 1 (the “best” health state) and scenario A has health state at level 2 (the moderate health state), A would still be chosen as the coefficient around the constant is greater than that of hl2 (the movement from health 1 to health 2).

Table 7.4. Probit model with constant term

<i>Attribute</i>	<i>Coefficient</i>	<i>Standard Error</i>
Health level 2	-0.079	0.033
Health 3	-0.462	0.057
Confidence 2	-0.081	0.024
Confidence 3	-0.468	0.038
GP access 2	-0.072	0.018
GP access 3	-0.254	0.032
Isolation 2	0.001	0.021
Isolation 3	-0.462	0.040
Constant	-0.120	0.017

where
health level 2 is movement from health state one to health state 2
health 3 is movement from health state one to health state 3
confidence 2 is movement from confidence level one to confidence level 2
confidence 3 is movement from confidence level one to confidence level 3
isolation 2 is movement from isolation level one to isolation level 2
isolation 3 is movement from isolation level one to isolation level 3
GP access 2 is movement from GP access level one to GP access level 2
GP access 3 is movement from GP access level one to GP access level 3

hl3/hl2 = 5.848
cl3/hl3 = 1.013

7.7.3 Random effects probit

Table 7.5 shows the results of the random effects probit with constant model. The results are very similar to those from the standard probit model in terms of the magnitude and direction of the coefficients for attributes. The similarity is not surprising as the random effects are attached to individuals and picking up omitted individual characteristics. However, as questionnaires were randomly allocated, there should be no correlation between individuals’ characteristics and

characteristics appearing in the regression model; therefore omitted variable bias should not occur. Again however, notice that the constant term is significant.

The health 3 to health 2 ratio in this model was 5.728, while the confidence 3 to health 3 ratio was 1.041, indicating the individuals considered a drop in confidence (from level 1 to 3) as marginally worse than a drop in health state (from 1 to 3), which again, is similar to the results of the standard probit model.

Table 7.5. Random effect probit model with constant term

Attribute	Coefficient	Standard Error
Health level 2	-0.081	0.033
Health 3	-0.464	0.057
Confidence 2	-0.107	0.024
Confidence 3	-0.483	0.038
GP access 2	-0.055	0.018
GP access 3	-0.284	0.032
Isolation 2	0.024	0.021
Isolation 3	-0.426	0.040
Constant	-0.148	0.017

$hl3/hl2 = 5.728$

$cl3/hl3 = 1.041$

7.7.4 Generalised Linear and Latent Mixed Models (GLLAMM)

As stated above, we would not expect a constant term to be significantly different from zero. However, in this instance, and in previous evaluations in the health economics literature, the constant term is significantly different from zero. In a labelled experiment, for example where scenario A is always based in a hospital setting, while B is in a home setting, this constant can be argued to reflect the preference for the “label” (in this example a preference for home over hospital or vice versa, other things equal). In an unlabelled experiment, the constant has no meaning and should be zero. The results of a GLLAMM model are presented in Table 7.6. A basic exploratory latent class model was estimated allowing for heterogeneous preferences across the population. The idea behind the latent class model is that there may be systematic groups who have a systematic set of different preferences. The objective of the latent class regression model is to not only estimate these preferences but also to estimate the proportion of these classes within the population and assign a probability of membership for each individual.

In this example, individuals are in two classes. The majority of individuals have a higher probability of being in class one (78%) rather than class two. It is noticeable that those in class one exhibits the same direction of values of attributes as the previous models (good health is preferred to bad and so on), though the size of these coefficients is larger than previously estimated. For this class of responder, the health 3 to health 2 ratio in this model was 3.285, while the confidence 3 to health 3 ratio was 0.793, indicating that the individuals in this class considered a drop in health state (from level 1 to 3) as considerably worse than a drop in confidence (from 1 to 3).

Table 7.6 GLLAMM model

Variable	<i>Latent Class 1</i>		<i>Latent Class 2</i>	
	<i>Coefficient</i>	<i>Std-errors</i>	<i>Coefficient</i>	<i>Std-errors</i>
Health level 2	-0.288	0.039	0.511	0.071
Health level 3	-0.946	0.058	0.651	0.094
Confidence level 2	-0.243	0.038	0.160	0.063
Confidence level 3	-0.750	0.049	-0.022	0.067
GP Access level 2	-0.017	0.035	-0.174	0.058
GP Access level 3	-0.240	0.044	-0.440	0.072
Isolation level 2	0.015	0.036	0.119	0.064
Isolation level 3	-0.649	0.040	-0.063	0.098
Constant	-0.219	0.028	-0.025	0.045
P(class)	0.784		0.216	

Hl3/hl2 (class 1)= 3.285
 cl3/hl3 (class 1)= 0.793
 hl3/hl2 (class 2)= 1.274
 cl3/hl3 (class 2)= -0.034

However, there are a sizable minority (approximately 22%) who have counter-intuitive values particularly for health state. This latent class appear to value decrements in health states whilst still valuing GP access (the other 2 attributes appear much less important to this group).

Health state 3, which is the ‘worst’ health state is preferred to health states 1 and 2. Although this raises some doubts about whether the responders who may be classified as class 2 responders have understood the exercise, it is this class that does not have a significant constant term. This raises the question of what to do with the preferences of this second class – do they genuinely represent odd

preferences or are they evidence that a subsection of responders has misunderstood the questionnaire?

In the event that it is the second reason, the effects of miss-responders in the homogeneous model are to weight the preferences of the population towards zero. Note how the ratio of moving down health states and confidence in the ability to manage condition increase relative to GP access (the level which is apparently unaffected by the latent class specification).

It should be noted that even though a significant term in a labelled experiment can be explained away by preferences for one or other label, this is not necessarily the only explanation. Heterogeneity of preferences could also explain part of this effect.

7.7.5 Random effects clustered on question number

A further possibility that we wished to explore was whether individuals adopted simple heuristics when options were closely matched. The notion being that if difficult trade-offs were required the respondent may adopt a simple heuristic of choosing right over left. This could be potentially picked up by running a probit model with random effects for each individual question. This will give each of the 38 questions their own specific constant term, a preference for B over A, after allowing for the differences between B and A. The results of this analysis are presented in Table 7.7.

Table 7.7 Random effects on question number

<i>Variable</i>	<i>Coefficient</i>	<i>Std-errors</i>
Health level 2	-0.087	0.046
Health level 3	-0.472	0.047
Confidence level 2	-0.109	0.046
Confidence level 3	-0.493	0.046
GP Access level 2	-0.052	0.044
GP Access level 3	-0.294	0.049
Isolation level 2	0.029	0.047
Isolation level 3	-0.414	0.046
Constant	-0.154	0.033
h12/h13 = 5.425		
cl3/h13 = 1.044		

The ratio of health level 2 to health level 3 is 5.425 and confidence level 3 to health state 3 is 1.044, indicating that a loss of confidence to manage their own condition is worse than a reduction in health state.

Unlike the random effects model which clustered results on individuals, the random effects model clustered on question number indicates that the random effects do indeed explain some of the variation (only a small amount) and that the standard deviation of the question specific constant terms differs from zero. Thus after allowing for what the options contain, some questions are more likely than others to have systematic preferences for B or A.

Given the preferences estimated, the model shows that questions 7 (59% chose B), 1 (57% chose B) and 21 (79% chose B) of the large questionnaire demonstrate an above average and increasing tendency for B to be chosen given the respective levels in A and B. Questions 13 (12% chose B) and 26 (81% chose B) of the large questionnaire and question 7 (76% chose B) of the short questionnaire have an increasing tendency to choose A despite the relative variables in both.

The 'middle' questions, i.e. those questions with average random effects are questions 4 (27%) and 6 (64%) of the short questionnaire and question 9 (13%) of the long questionnaire.

7.7.6 Sensitivity analyses

Two questionnaires were administered and the majority of the analysis above is based on the merged data from both questionnaires. However, a random effects probit model was performed on each questionnaire to identify potential differences in responses. The results of this analysis are presented in Table 7.8. Though the direction of the results is the same for all attributes, the magnitude of the effects is very different for several of the attributes. The interaction terms described above may have some impact on these findings, though other effects, such as learning effects may also be responsible.

Table 7.8 Responses to short and long questionnaires

	<i>Short questionnaire coefficient</i>	<i>Short SE</i>	<i>Long questionnaire coefficient</i>	<i>Long SE</i>
Health level 2	-0.234	0.058	-0.064	0.048
Health 3	-0.064	0.101	-0.429	0.082
Confidence 2	-0.186	0.047	-0.081	0.033
Confidence 3	-0.762	0.079	-0.434	0.054
GP access 2	-0.021	0.042	-0.068	0.024
GP access 3	-0.475	0.072	-0.277	0.043
Isolation 2	0.115	0.050	0.006	0.027
Isolation 3	-0.118	0.049	-0.475	0.059

The impact of including “inconsistent” responses is presented in Table 7.9. Most of the attributes, with the exception of GP access level 2, maintain the same direction. The GP level 2 attribute becomes positive, implying that responders prefer to see their GP in one week’s time rather than in one day’s time. The magnitude of the coefficients around the other attributes is altered when compared with the “consistent” responses analysis, though the relative impact of self-efficacy compared to health status remains similar.

Table 7.9. Inclusion of “inconsistent” responses

<i>Attribute</i>	<i>Coefficient</i>	<i>Standard Error</i>
Health level 2	-0.080	0.026
Health 3	-0.286	0.043
Confidence 2	0.097	0.026
Confidence 3	-0.232	0.037
GP access 2	0.021	0.020
GP access 3	-0.114	0.029
Isolation 2	0.122	0.022
Isolation 3	-0.180	0.040
Constant	-0.150	0.016

7.7.7 Impact of inclusion of self-efficacy in the CEAC

The additional impact of self-efficacy on QALYs from the EPP depends on the numbers moving from one state of self-efficacy (low, moderate or very confident). In this crude analysis, there is no QALY benefit associated with a movement within category (for example a movement from 8 to 10 on the self-efficacy question), nor is there disbenefit associated with drops within categories. Individuals only receive additional QALY benefit (disbenefit) if they improve (reduce) their category. The numbers moving between categories by group (intervention vs control) is displayed in Table 7.10. More individuals in the intervention group than in the control group

improved from the worst self-efficacy state to the best, while more in the control group recorded disbenefits in that they moved to a lower category.

Table 7.10. Numbers moving between self-efficacy category by group.

	<i>Intervention group</i>	<i>Control Group</i>
Improved from worst self-efficacy to best	22	10
Worst to moderate	18	28
Moderate to best	46	42
Best to worst	1	4
Moderate to worst	15	23
Best to moderate	103	109

Figure 7.2 illustrates the effect of including self-efficacy as an additional outcome on the results of the EPP national evaluation presented in Chapter 4. Interventions that impact positively on self-efficacy will, other things equal, be more cost-effective when these outcomes are included (so long as there is a positive relationship between HRQoL and the chosen outcome). In this instance, the EPP intervention originally generated 0.020 QALYs compared with the control group and had a probability of 94% of being cost-effective at a threshold value of £20,000 for a QALY. The inclusion of the additional impact of self-efficacy increases the QALYs gained to 0.025 and increases the probability of the intervention being cost-effective to 96% at a threshold value of £20,000 for a QALY.

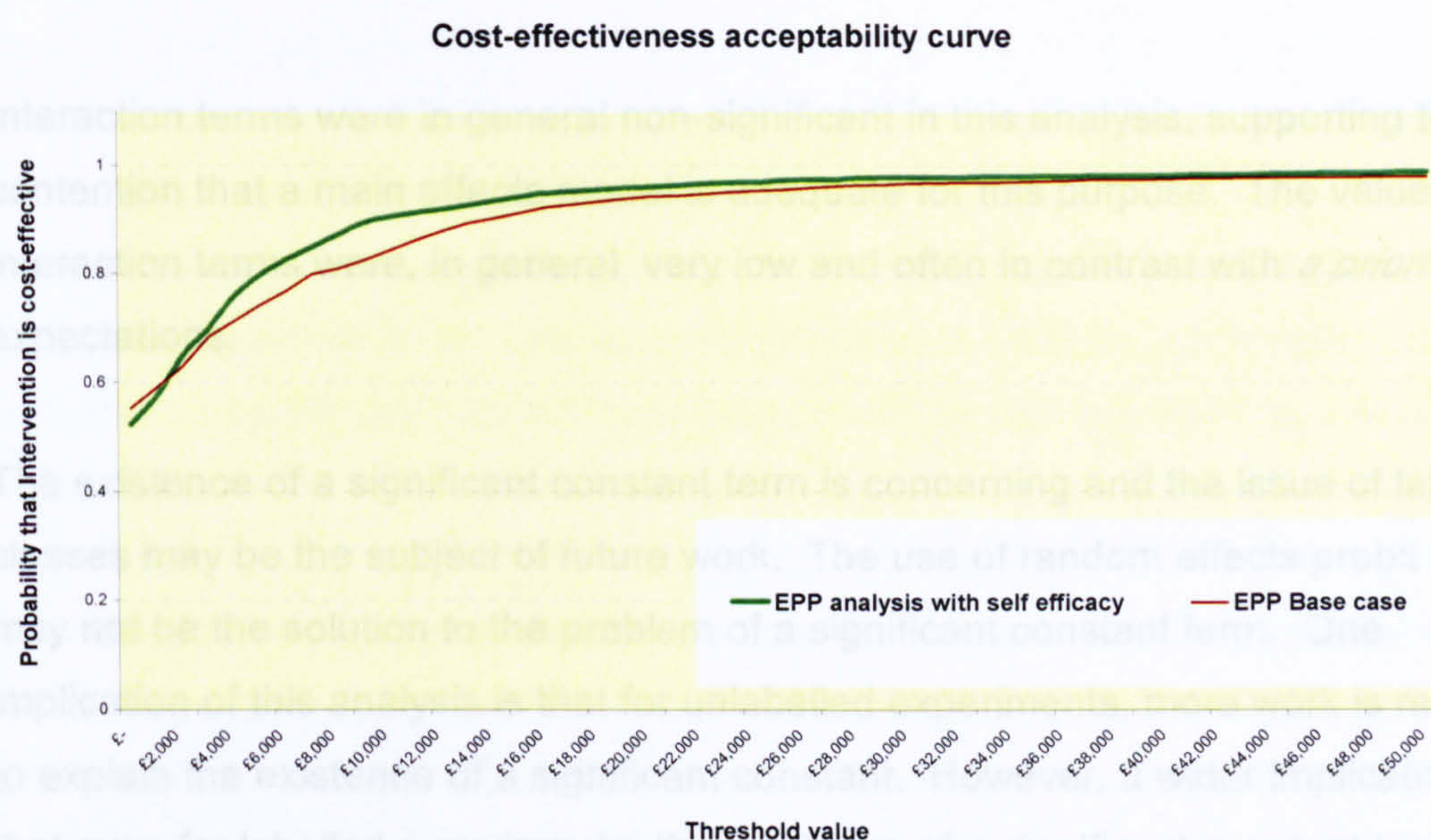


Figure 7.2. Impact of including self-efficacy on probability of EPP being cost-effective

7.8 Discussion

A response rate to a postal questionnaire of over 70% indicated that patients were willing to complete the DCE. The length, and therefore the cognitive burden of the questionnaire, had little impact on the response rate. Whichever model was chosen, the responses to both questionnaires were largely consistent with *a priori* expectations. The results of this DCE indicate that self-efficacy is an important outcome measure, and patients were willing to trade decrements in HRQoL for improvements in self-efficacy, and that this rate of trade-off was broadly similar across all models. Using a crude additive model, these estimates of the importance of self-efficacy can be included in the cost-effectiveness analysis. Their inclusion has no impact on the costs, but increases the QALYs gained from 0.020 to 0.025 per person per year. These estimates are derived from the valuation of movements between levels of confidence and the numbers making these moves. Thus, a movement from not at all confident to totally confident is approximately the same as a movement from health state 3 to health state 1. As the latter movement is valued at approximately 0.25 QALY, so the movement from not at all confident to totally confident is valued similarly. However, from Table 7.10, only a minority of patients make this change, and the impact on results is not large. The associated CEAC

shows that the probability of the EPP being cost-effective increases from 94% to 96% at £20,000 per QALY.

Interaction terms were in general non-significant in this analysis, supporting the contention that a main effects model is adequate for this purpose. The values of interaction terms were, in general, very low and often in contrast with *a priori* expectations.

The existence of a significant constant term is concerning and the issue of latent classes may be the subject of future work. The use of random effects probit models may not be the solution to the problem of a significant constant term. One implication of this analysis is that for unlabelled experiments, more work is required to explain the existence of a significant constant. However, a wider implication is that even for labelled experiments, the existence of a significant constant term is worrying. While this has previously been treated as simply a preference for the label (e.g. home over hospital), it is possible that this is not the only explanation.

A further concern is the number of responders who exhibited either a tendency to give inconsistent responses and/or responses that were counter-intuitive (such as valuing “worse” health states higher than “better” health states). Around 27% of responders gave inconsistent responses. Of the remainder, 22% gave counter-intuitive values for health states. While the direction and magnitude of the coefficients around the attributes lend weight to the assertion that the DCE was broadly understood and answered correctly, there remains a suspicion that a relatively large number of responders did not understand the exercise. Different DCEs use alternative methods to assess rationality and/or consistency. San Miguel²⁶⁹ found that depending on the definitions employed, between 1 and 35% of patients could be described as irrational, in that they failed to select a dominant option.

Typically, economic evaluations employ the QALY as a measure of HRQoL. QALYs generated from the EQ-5D instrument do not include self-efficacy explicitly and thus QALYs generated in this manner may omit the importance of self-efficacy. Including self-efficacy in the CEA would increase the probability, *ceteris paribus*, of

interventions that improve self-efficacy being cost-effective. DCEs are one methodology that allows the inclusion in the QALY.

However, there are a number of caveats. Firstly, it is not clear whether self-efficacy is a health outcome; if it is not then should it be included in the objective function of a decision maker working within a budget constrained health system? Should these outcomes be considered? Decision makers need to be clear *a priori* about the maximand and not use vague statements about other factors that may be taken into account.

Secondly, self-efficacy may simply be a measure of process in that patients want self-efficacy as they perceive that this will improve their long-term health status. The study surveyed people with chronic conditions and it is likely that these valuations of self-efficacy will be different from those of the general public. Thus, a limitation of this study is that these patients' values may overstate the importance of self-efficacy, and thus make interventions that improve self-efficacy appear more cost-effective. This is a problem whenever stated preference techniques are used in non market situations. If we are interested in getting patient valuations of outcomes, then we need to be aware that patients may give higher values than they would truly be willing to pay.²⁷⁰

Thirdly, it is conceivable that changes in self-efficacy (or part of those changes) are already incorporated into the QALY, through one or more of the dimensions of the chosen instrument.

Though self-efficacy has been identified as important to patients, it is likely that in different patient groups other outcomes would be valued and traded for HRQoL. These "important" outcomes should be identified before commencing the study and appropriate techniques should be used to ascertain the rate of substitution between these outcomes and HRQoL.

Finally, it is acknowledged that the use of DCEs may not force responders to focus on the real opportunity cost sacrifice to health by presenting a direct trade-off between health outcome and self-efficacy. This is likely to result in a greater chance

of over-stating the relative value of the self-efficacy. Cookson demonstrated that individuals expressed larger relativities when trade-offs were expressed in monetary terms rather than lives saved – similarly in this instance trading-off self-efficacy may be more palatable and therefore result in higher valuations than if individuals were asked to reduce (for example) life expectancy.²⁷¹

So what implications does the analysis in this chapter have for economic analysis on interventions to support self care? The arguments around the perspective of studies and whose values should be used were discussed in Chapter 2. At present, government appointed decision making bodies, such as NICE, use general population values for health states. Some authors argue for inclusion of patient valuation of health states (rather than valuation by the general population),²⁷² while others encourage the assessment of other outcomes that are important to patients.²²¹ While the patients' role in identifying and valuing outcomes may become more important as we move to a more "patient centred" service, the choice of outcomes and the methodology used to derive their values needs to be explicitly stated and the caveats and assumptions implied by this choice should be emphasised.

7.9 Conclusions

The above analysis has demonstrated that individuals are prepared to answer questionnaires with a large cognitive burden. However, a large minority may have difficulty understanding the questionnaire and this is not dependent on the length of questionnaire.

Self-efficacy is valued in this patient population, and can be incorporated into the QALY estimates relatively easily, though caution should be exercised in doing so. Other outcomes were also valued by patients and in principle, these could be incorporated into the analysis also.

Chapter 8. Discussion

Self care and support for self care have become important policy objectives. This thesis has examined whether this objective is justified on cost-effectiveness grounds.

8.1 Summary of individual chapters

Chapter 1 outlined the decision problem of whether interventions to support self care provide additional benefits that are worth paying for in the budget constrained system of the UK NHS. It was argued that economic evaluation provides an explicit methodology for addressing this decision problem, by establishing the relative cost-effectiveness of interventions compared with suitable alternatives.

However, economic evaluation encompasses a range of opinions and is not free from controversy. The choice of paradigm can affect the design conduct, analysis, results and recommendations of an economic evaluation. It is important that the analyst is explicit in the choice of paradigms and the assumptions that these choices imply.

In Chapter 2, common controversies within economic evaluation were examined. Historically, most economic evaluations have been conducted within or alongside clinical trials (typically RCTs), and these “single trial based analyses” dominated the literature. More recently, it has become acknowledged that there are limitations to this approach. Single trial based economic evaluations (and indeed single clinical trial analyses) suffer from a number of defects that relate to the exclusion of relevant information (all four of the deficiencies noted in Chapter 2 can be simplified to “ignoring information that may be relevant to the decision problem”). Nevertheless, there are occasions when a single trial based analysis may be the most appropriate evidence to inform the decision problem, such as instances where the data come from a very different population or the other studies were poorly conducted.

The distinction between the Bayesian and frequentist schools was also discussed in Chapter 2. Bayesians consider probabilities to be reflections of a degree of belief and that this probability can be amended as more evidence becomes available. This enables Bayesians to make statements about the “probability” of an

intervention being cost-effective that are partly dependent on “prior beliefs”. Frequentists cannot make such statements as probabilities are simply a function of the frequency of events in a much repeated experiment. It was argued that the Bayesian approach is more useful for informing the decision problem but that care should be taken around the elicitation and influence of the priors.

Chapter 2 also provided a detailed explanation of the key differences between the welfarist and extra-welfarist perspectives. It was asserted that the distinctions between the two schools are often confused, and there are few definitive guides to welfarism and extra-welfarism. Welfare (or welfarist) economics assumes that individuals maximise their utility, that individuals are the best judges of their own utility (individual or consumer sovereignty), that utility is a consequence of behaviour and that utility information *only* is relevant in making decisions about what is best. The last named is frequently termed “welfarism”. The practical problems of collecting this utility information, plus the restrictions placed on the sources of these valuations (must be the individual), the weights attached to them (it is unclear whether weights are allowed) and the lack of interpersonal comparisons mean that for the purposes of this thesis, the extra-welfarist approach was deemed appropriate. Extra-welfarism allows information outside individual utilities (hence the “extra”). Paradoxically, extra-welfarist analyses have tended to adopt narrower perspectives by measuring a single health related outcome. Some analyses consider overall health by using measures such as the QALY, while others use much narrower measures such as reduction in blood pressure.

It is important to stress that the choice of paradigms may impact on the cost-effectiveness study and how the results of these analyses inform the decision problem. The published literature on self care support interventions is examined in a systematic review presented in Chapter 3. This chapter presented the first review of economic evaluations of self care support interventions. The chapter described the paradigms which have been used in previously published economic evaluations and also to assess whether any lessons could be learned for the design, conduct and analysis of future evaluations.

Studies identified for inclusion in the systematic review presented in Chapter 3 included numerous disparate interventions for a range of conditions, in a range of settings and geographical locations. The results of the systematic review showed that all the identified studies were based in the extra-welfarist paradigm, with patients' health outcomes as the measure of interest. Studies were rarely synthesised with other relevant data, either in a narrative or quantitative manner, to address the decision problem more comprehensively. The majority of studies were poorly conducted and/or analysed, though the conclusions were mainly in favour of the intervention to support self care. The studies in general were of limited use in addressing the decision problem. Disease specific outcome measures, or behavioural/psychological outcome measures, were frequently employed without considering the value of these outcomes. Thus comparisons of cost-effectiveness across conditions were problematic. The chapter concluded with a call for the use of generic outcome measures in economic evaluations to allow decision makers to compare the cost-effectiveness of interventions across conditions.

Chapter 4 presented an original single trial based economic evaluation using a generic instrument as a measure of outcome, the QALY. This economic evaluation, which is currently in press, was conducted alongside the national evaluation of the Expert Patients Programme (EPP), and represents the first economic evaluation of the EPP in the UK.^{66, 162} This study is set in the extra-welfarist paradigm and employs Bayesian techniques to assess the probability that the EPP intervention is cost-effective at a variety of threshold values of a QALY. The main analysis showed that the EPP based solely on this study was approximately 94% likely to be cost-effective at a threshold value of a QALY of £20,000 when compared with a waiting list control. There were a number of limitations to this study. Firstly, the comparator was in effect a usual care alternative. Individuals in the control group would receive the intervention at the end of the six month period, but until then received the care they had previously received. This is a limitation because it is plausible in this study that the effects of being in a group of individuals with similar conditions may itself have improved patient outcomes. It is difficult to separate the effect of the lay led delivery of advice from the effect of being in a group of similar individuals. In addition, the follow-up period is relatively short. It is possible that in the longer term the benefits of the EPP may be maintained or may dwindle and the

long-term cost-effectiveness is therefore debatable. Extrapolating the results to one year in a simplistic sensitivity analysis showed an improved probability of EPP being cost-effective at commonly used threshold values of a QALY, but is based on the simplifying assumption that costs and outcomes were equally distributed over the six month period. Other distributions are plausible and these would affect the conclusions.

While these are important limitations, the main criticism of the evaluation presented in Chapter 4, is that it may not consider all information of relevance to the decision problem. At the time of writing, there was one other RCT of the EPP based in the UK,⁸³ and several other trials and experiments based outside the UK using interventions identical or similar to that used in the EPP.^{71, 75, 76, 78-82} This potential criticism of single trial based analyses was addressed in Chapters 5 and 6. Expanding the evidence base initially to two RCTs based in the UK, as in Chapter 5, could clearly impact on the decision problem. Using Individual Patient Data (IPD) from both these trials enabling a more complete meta-analysis including the impact of individual characteristics and avoiding the pitfalls commonly associated with meta-analysis of aggregate data. Though these two trials considered the same intervention and comparison, they were very different in terms of the populations recruited and the results obtained. The national evaluation, with a target population designed to be representative of the UK general population demonstrated an improved QALY profile at slightly reduced cost. The study of EPP in a UK based Bangladeshi population showed different direction of results. QALY profiles were negative (after the appropriate adjustment for baseline EQ-5D score) with an increase in costs. While the direction is different for both costs and effects, the incremental costs and incremental QALYs reported in both studies were small.

The two studies were of comparable size (629 in the national evaluation, 476 in the Griffiths study), and were both judged to be of equal weight in the analysis. The cost of the intervention was estimated as a weighted mean of the intervention costs in the two trials (weighted by sample size). Using these estimates of costs and effects and the log normal distribution that best fitted the data, the EPP intervention generated an ICER of around £30,000 and a corresponding probability of being cost-effective of 50% (at the £30,000 threshold value of a QALY). At lower

threshold values of a QALY, say £20,000, the EPP intervention is unlikely (at only 20%) to be cost-effective.

The QALY difference between the intervention and control arms becomes very small when performing this synthesis. With such small values for the denominator (incremental QALYs), the ICER becomes very sensitive to small changes in costs. The cost of the intervention in the national evaluation was estimated at £250, and if this were replicated, the intervention is unlikely to be cost-effective (with an ICER of £41,500), while using the estimate from the Griffiths study yields an ICER of £20,333 which is potentially cost-effective. A limitation of the synthesis, and the two trials that make up the synthesis, is the lack of a measure of uncertainty around the cost of the intervention. For the purposes of the preceding analysis, the intervention cost is considered as a point estimate. It is relatively common in economic evaluation to consider unit costs as point estimates and subject them to sensitivity analysis to establish their impact on results. However, given the sensitivity of results to this variable, future work should consider the possibility of establishing the uncertainty around the point estimate of the unit cost of the intervention.

As well as the limitation of the use of a point estimate for the unit cost of the intervention, the analysis presented in Chapter 5 could also be criticised for not incorporating all relevant information or evidence. Though both RCTs based in the UK using the EPP intervention on a range of individuals with a range of chronic conditions were included, other evidence from outside the UK could be useful to inform the decision problem.

Since the other studies did not use QALYs as the outcome measure, but rather surrogate or intermediate outcomes such as self-efficacy, it was necessary to examine methods to incorporate evidence from outside the UK that used a surrogate measure of outcome. This is considered in Chapter 6. This chapter used systematic searching to identify evidence of the impact of the EPP intervention on costs and outcomes. No other trials were identified that measured the outcomes in terms of QALYs. However, a further eight trials were identified that presented data (in aggregate form) of an alternative outcome measure, self-efficacy. For the purposes of this analysis, self-efficacy was interpreted as a predictor of health state.

Data recorded in the two UK trials of EPP described in Chapter 5 included costs, QALYs and self-efficacy. Thus a link between the EPP intervention and costs and QALYs could be estimated indirectly using self-efficacy as a surrogate or intermediate outcome. This “indirect” link was employed in Chapter 6 to assess the impact of introducing the eight additional non-UK studies to inform the decision problem. Utilising this link enables three models to be developed. The “direct” model where the indirect evidence is not incorporated into the analysis is considered. The “indirect” model is then examined where only evidence utilising the link between treatment and Costs/QALYs *via* self-efficacy is considered. Finally, the “mixed” or “joint” model, where both forms of evidence are admitted, is analysed. Each of the eight studies included in the meta-analysis of IPD and Aggregate Data (AD) presented data on the effect of the intervention on self-efficacy. Each showed a positive relationship between the intervention and the level of self-efficacy (that is self-efficacy improved in those receiving the intervention). Given this positive relationship incorporation of these studies would improve the estimates of the effectiveness of the intervention as long as there was a positive relationship between self-efficacy and QALYs. While this relationship, estimated from the two UK studies, was positive it was not particularly strong. Thus although the estimate of effectiveness increased with the inclusion of this additional evidence, the impact on the ICER was limited. Interestingly though, because of the small differences in costs and QALYs described earlier, even small changes in either costs or effects can influence the decision. This is the case in this instance where the decision may be influenced as the ICER falls from around £30,000 per QALY to around £23,000.

While these differences may be important, the difference between the indirect and other models is perhaps of more concern. The indirect model utilised only evidence linking the treatment to cost and QALY changes via the surrogate outcome self-efficacy. The decision not to implement the intervention would be made with virtual certainty at commonly used threshold values of a QALY, while further research would be deemed unnecessary (again with virtually no benefit to conducting the research). This is because using this model, we are very confident that the ICER lies above reasonable threshold values of a QALY. This is in contrast with the largely equivocal results from the direct analysis and the positive results from the single trial based analysis presented in Chapter 4.

So, the decision is sensitive to the choice of model. Which model is “correct”? This depends on what evidence is available and what evidence is considered relevant. The indirect model described above may be a practical choice if there is no other evidence available and a decision has to be made. However, this decision should be made in the knowledge that the structural assumptions of the model have not been tested in any way and that uncertainty in the adoption decision and the decision to conduct future research may be understated.

Where more data are available in the form of IPD from randomised controlled trials, these should be used to inform the decisions. However, this does not sidestep the issue of relevance. Is an RCT in the US relevant to a UK decision maker? It may be argued that it is somewhat relevant, but not as relevant as a UK trial in a representative population. So what weight should it be given? Half? Quarter? Who decides how relevant the trial is and on what basis? Is it the responsibility of the analyst to present analyses that allow the decision maker to select the appropriate model given his/her definition of “relevant”, or should the analyst be more affirmative? Inevitably, the analyst will be required to make some choices in the representation of results and recommendations, otherwise decision makers will be faced with unwieldy analyses. Nevertheless, as argued in Chapter 2, the democratic process is deemed an appropriate mechanism to appoint legitimate decision makers. While they cannot be expected to consider every possible alternative, the analyst should provide them with scenarios examining the influence of pivotal variables or data. Thus the analyst has influence over the analyses presented but the decision as to which analysis influences policy is left to the legitimate decision maker.

In this instance, a reasonable question from a policy maker would be “given all these potentially different models, datasets and assumptions, what is the recommendation for policy based on existing evidence?”. As highlighted throughout this thesis, there is no definitive answer to this question. Intuitively though, it is appealing that the national evaluation with a representative sample of the population should present the best evidence of the cost-effectiveness of the EPP in a UK setting. Where the comparison is treatment as usual, it is reasonable to conclude that the EPP is a

cost-effective use of resources at commonly employed threshold values of a QALY. The inclusion of other comparators may alter this conclusion

The discussion above has centred on the relevance of data from outside the UK using alternative measures of outcome. The implicit assumption is that the only outcome of interest is “health” and that this is adequately captured by the QALY. However, it is plausible that there are other outcomes that may be considered relevant to the decision problem. Indeed, as argued in Chapter 2, the extra-welfarist paradigm has its foundations in a movement away from the narrow concept of utility. It may be that decision makers are interested in either a broader concept of health (including some psychological outcomes), or that we wish to go beyond health and introduce measures such as isolation. In the field of self care support, the concept of self-efficacy has attracted much attention since the 1970’s. However, it is not clear from the published literature how a decision maker could use this outcome to inform the decision problem as it is not widely used outside the self care support field and importantly, has not been subjected to valuation. Chapter 7 addresses this problem by using a discrete choice experiment to value self-efficacy compared to health, as measured by the QALY, and other outcomes that were assessed as being important to patients with chronic conditions. Individuals who completed the questionnaire valued self-efficacy and in principle, these valuations can be incorporated into cost-effectiveness analysis. This is the first attempt to value this outcome measure and should allow a more useful assessment of treatments where this is the outcome measure of choice.

8.2 What could have been done differently?

Most of the evaluations in the published literature consider the comparison of an intervention with treatment as usual. In many cases, this is the comparison of interest and directly addresses the decision problem. However, in other cases, particularly the use of group therapy for chronic conditions, the decision problem may be wider and require information on a range of potential comparators. For example, the Chronic Disease Self Management Programme may be shown to be cost-effective when compared to treatment as usual, but how does it compare with other group based interventions that may be available for the same population? The

evaluations of CDSMP presented in the literature and in Chapter 4 have, largely for practical reasons, limited the comparison to treatment as usual. Thus the analyses inform a more limited decision problem. However, the analyses presented in Chapter 6, provides a framework for including additional comparators. This is discussed in the section below on recommendations for future research.

The studies included in the analyses presented in Chapters 3-6 had a limited follow-up period. Again, this is likely to be for practical reasons, but the longer term impact of these interventions remains uncertain. While analyses in Chapters 4 and 5 extended the follow-up period to a longer time frame, this was based on reasonable, but unsubstantiated, assumptions. It is not certain when the additional costs and effects of the interventions occur. This could certainly impact on the cost-effectiveness of the interventions not only for extending the time frame but also for the existing analysis. For example, QALYs are calculated as a product of time in a given state and the “health” of that state. Typically, health is measured using a generic health related quality of life instrument such as EQ-5D. The instrument is usually employed at baseline and follow-up(s), and a linear relationship between EQ-5D scores at these time points is invariably assumed.⁴³ In the absence of evidence to the contrary this may be a reasonable assumption. However, in reality, the benefit of the intervention may not be distributed evenly and may, for example, occur at the beginning of the period with the effect dissipating over time. Thus the effectiveness of the intervention would be understated, if the short-run benefits are not acknowledged as they fall between follow-up points. Similarly, the intervention may have little impact on health until it is completed, that is, most of the benefit will occur at the end of the intervention (near the follow-up measurement). Linear interpolation of the EQ-5D scores will overstate the benefit of the intervention. Given the small beneficial effects of the intervention evaluated in Chapters 4-6, the distribution of effects (and costs) could be pivotal.

The search strategies developed in Chapters 3 and 6 to review existing economic evaluations and to populate the models respectively could have been susceptible to publication bias. In particular, the results of the searches in Chapter 6 only identified studies that reported a positive influence of the intervention on the (surrogate) outcome. The majority of studies identified in the review of economic evaluations

considered in Chapter 3 and the review of studies to populate the decision model in Chapter 6, were of low quality. The conduct, design, analysis and reporting of results varied considerably between studies. The analyses in this thesis made no attempt to weight studies according to quality as it is unclear what would be appropriate weights to give to studies of different quality. Nevertheless, a simple sensitivity analysis excluding poor study designs showed little impact on the ICER, and no effect on the likely results, conclusions and recommendations of the analysis.

The analyses used in Chapters 4, 5 and 6 employed a generic outcome measure of health, the QALY, which it has been argued is most useful in addressing the decision problem. The QALY in this instance was calculated using the EQ-5D instrument whose scores were then “valued” using a sample of the UK general public. This is considered the appropriate methodology for current decision makers,¹² but is not universally accepted.^{272, 273} There is concern that the QALY does not capture all outcomes of interest (whether they are health outcomes or not) and that patients’ may give the most reasonable values for the health states of which they have experience. In Chapter 7, both of these issues were examined using a discrete choice experiment. The sample population consisted of a convenience sample of patients who had been offered the EPP intervention and were asked to value self-efficacy compared to health related quality of life and other important outcomes. This patient group appeared to value the concept of self-efficacy, and were prepared to trade-off improvements in self-efficacy for decrements in health related quality of life. However, while these results suggest that self-efficacy has some value, the fact that the patient population had been trained in the concept during the EPP, means that the estimate of the relative worth of self-efficacy may be overstated compared to the value of self-efficacy in a more general chronically ill population (those who had not been on the EPP). In order to estimate more plausibly the value of self-efficacy, it would be appropriate to conduct the experiment in a less biased group.

The analyses presented in Chapters 4, 5 and 6 used a point estimate of the cost of the intervention. Uncertainty around unit cost data is not always considered in cost-effectiveness analyses as unit costs tend to be estimated from tariffs or previous

publications that only present point estimates. However, in this instance, the results of the analyses are particularly sensitive to the cost of the intervention. While standard univariate sensitivity analysis has been employed, obtaining a distribution around the point estimate of the cost of the intervention would be preferable. This should be achievable given that number of attendees of each group, for example, could be easily recorded.

8.3 Recommendations for future research

The limitations to the analyses presented above form the basis of the recommendations for future research.

The cost-effectiveness of interventions to support self care is questionable in general. In particular, the cost-effectiveness of the EPP depends on the assumptions we make about the relevance of evidence. However, the analyses presented above compare the EPP to treatment as usual rather than an active comparator such as another group therapy. The latter comparison is more useful in addressing the decision problem as it informs not only whether a specific intervention is cost-effective compared with treatment as usual, but also which intervention(s) are likely to be most cost-effective compared with each other. While this thesis has not sought to introduce a network of evidence comparing many alternatives, the framework presented in Chapter 6 allows just such analyses to be performed. This is a key area for future research in this field.

The analyses performed in this thesis have been conducted with the benefit of two datasets with individual patient data and a sizable published literature. It has been shown that the single trial based analysis presented in Chapter 4 may generate different recommendations for the cost-effectiveness of the EPP than do the analyses presented in Chapters 5 and 6. The latter chapters introduced other potentially relevant evidence, and it is this additional evidence which potentially alters the results and conclusions. It is necessary to state at the outset what evidence, and sources of evidence, will be considered potentially relevant to the decision problem. The direct, indirect and mixed models employ differing assumptions and mechanisms. It is unclear which model is appropriate for decision

making purposes, where more than one form of data is available. More research is required to assess whether the unexplained variation in the direct model reflects genuine variation or is simply noise that we can safely ignore.

Future economic evaluations of these interventions should also collect information that quantifies the uncertainty around the distribution of the cost of the intervention. This will allow a more accurate representation of the uncertainty around the decision.

8.4 Conclusion

This thesis has developed the existing evidence to inform the decision problem of whether self care support interventions in general, and the EPP in particular, are cost-effective. This thesis presents the first (and to date, the only) single trial based economic evaluation of the EPP in the UK. The evidence base is expanded to include other trial data and published literature using techniques that have not previously been reported. The findings are equivocal. Whether these interventions are cost-effective or not depends on

- choice of paradigm
- concept of relevance
- distributional assumptions
- choice of comparison

Different concepts will generate varied results with alternative recommendations for decision makers. The concern with the use of single trial based analyses as a vehicle for informing the decision problem is illustrated by comparing the results of this analysis with analyses using more evidence. However, the single trial based analysis presented in this thesis presents the most relevant evidence and the results suggest that the EPP is a cost-effective use of resources when the comparator is treatment as usual. Whichever model is chosen as the basis for the assessment of cost-effectiveness, analysts need to be explicit about the assumptions required by the analysis. Failure to do so may lead to an underestimation of the uncertainty around the decision to adopt the technology as well as an understatement of the value of conducting future research.

References

1. Epping-Jordan J, Bengoa R, Kwar R, Sabaté E. The challenge of chronic conditions: WHO responds. *BMJ*2001;323:947.
2. Walker A, O'Brien M, Traynor J, Fox K, Goddard E, Foster K. *Living in Britain. Results from the 2001 General Household Survey*. 2002 [cited 2007 14 Nov]. Available from: <http://www.statistics.gov.uk/lib2001/Section3476.html>
3. Department of Health. *Self care - a real choice: Self care support - a practical option*. London: Department of Health; 2005.
4. Wagner E, Groves T. Care for chronic diseases. *BMJ*2002;325:913-4.
5. Wanless D. Securing our Future Health: Taking a Long-Term View. 2002.
6. Griffiths C, Foster G, Ramsay J, Eldridge S, Taylor S. How effective are expert patient (lay led) education programmes for chronic disease? *BMJ* 2007;334:1254-1256.
7. Richardson G, Gravelle H, Weatherly H, Ritchie G. Cost-effectiveness of interventions to support self-care: a systematic review. *Int J Technol Assess Health Care* 2005;21:423-32.
8. Department of Health. *The NHS Plan: a plan for investment, a plan for reform*. London: Department of Health; 2000.
9. Donaldson L. The expert patient: a new approach to chronic disease management for the 21st century. *BMJ*2003;326:1279-80.
10. Donaldson L. Expert patients usher in a new era of opportunity for the NHS. *BMJ*2003;326:1279-80.
11. Sullivan SD, Lyles A, Luce B, Grigar J. AMCP guidance for submission of clinical and economic evaluations data to support formulary listing in US health plans and pharmacy benefits management organisations. *J Manag Care Pharm* 2001;7:272-82.
12. National Institute for Clinical Excellence. *Guide to the methods of technology appraisal*. London: NICE; 2004.
13. Ontario Drug Benefits Program. *Ontario guidelines for drug submission and evaluation*. 2000 [cited 2007 14 Nov]. Available from: www.health.gov.on.ca/english/providers/pub/drugs/dsguide/dsguide_mn.html
14. Commonwealth Department of Health Housing and Community Services. Guidelines for the pharmaceutical industry on preparation of submissions to the Pharmaceutical Benefits Advisory Committee. 2002 [cited 2007 14 Nov]. Available from: <http://www.healthconnect.gov.au/internet/wcms/publishing.nsf/Content/health-pbs-general-pubs-guidelines-index.htm>
15. Drummond M. Should commercial-in-confidence data be used by decision makers when making assessments of cost effectiveness? *Appl Health Econ Health Policy*2002;1:53-4.
16. Drummond M. Will there ever be a European drug pricing and reimbursement agency? *Eur J Health Econ* 2003;4:67-9.
17. Brouwer W, Culyer A, van Exel N, Rutten F. Welfarism vs extra-welfarism. *J Health Econ* 2007;in press.
18. Sen A. Social choice theory: a re-examination. *Econometrica* 1977;45:53-89.

19. Hurley J. An overview of the normative economics of the health sector. In: Culyer A, Newhouse J, editors. *Handbook of Health Economics*. Amsterdam: Elsevier; 2000. p. 55-118.
20. Sen A. Capability and well-being. In: Nussbaum M, Sen A, editors. *The Quality of Life*. Oxford: Clarendon Press; 1993. p. 62-67.
21. Culyer AJ. The normative economics of health care finance and provision. In: McGuire A, Fenn P, Mayhew K, editors. *Providing health care: the economics of alternative systems of finance and delivery*. Oxford: Oxford University Press; 1991. p. 65-98.
22. Sugden R, Williams A. *The principles of practical cost-benefit analysis*. Oxford: Oxford University Press; 1978.
23. Farrow S. Environmental equity and sustainability: rejecting the Kaldor-Hicks criteria. *Ecol Econ* 1998;**27**:183-8.
24. Zerbe R, Bauman OY, Finkle A. An aggregate measure for benefit-cost analysis. *Ecol Econ* 2006;**58**:449-61.
25. McGuire A, Henderson J, Mooney G. *The economics of health care: an introductory text*. London: Routledge; 1998.
26. Mooney G. Judging goodness must come before judging quality - But what is the good of health care? *Int J Qual Health Care* 2000;**12**:389-94.
27. Sen A. *Commodities and capabilities*. Oxford: Oxford University Press; 1985.
28. Sen A. *On economic inequality*. Oxford: Clarendon Press; 1973.
29. Sen A. The impossibility of a Paretian liberal. *J Polit Econ* 1970;**78**:152-7.
30. Arrow K. *Social choice and individual values*. New York: John Wiley and Sons; 1951.
31. Sen AK. Personal utilities and public judgments: or what's wrong with welfare economics. *Econ J* 1979;**89**:537-58.
32. Ng Y. Welfarism: a defence against Sen's attack. *Econ J* 1981;**91**:527-30.
33. Scitovsky T. *The joyless economy: an inquiry into human satisfaction and consumer dissatisfaction*. New York: Oxford University Press; 1976.
34. Mishan E. *Economic efficiency and social welfare : selected essays on fundamental aspects of the economic theory of social welfare*. London: Allen & Unwin; 1981.
35. Varian H. *Microeconomic analysis 3rd edition*. New York: Norton; 1992.
36. Birch S, Donaldson C. Valuing the benefits and costs of health care programmes: where's the 'extra' in extra-welfarism? *Soc Sci Med* 2003;**56**:1121-33.
37. Claxton K, Sculpher M, Culyer A. Mark versus Luke? Appropriate methods for the evaluation of public health interventions. *Health Econ* 2007;(in press).
38. Sculpher M, Claxton K, Akehurst R. It's just evaluation for decision making: recent developments in, and challenges, for cost-effectiveness research. In: Smith P, Ginnelly L, Sculpher M, editors. *Health Policy and Economics: Opportunities and Challenges*. Maidenhead: Oxford University Press; 2005. p. 8-33.
39. Lipsey R, Lancaster K. The general theory of second best. *Rev Econ Stud* 1956-7;**24**:11-32.
40. Richardson J. *What can health economics offer?(CHPE Working Paper 130)*. Melbourne: Centre for Health Program Evaluation; 2002.
41. Sen A. Why health equity? *Health Econ* 2002;**11**:659-66.
42. Weinstein M, Stason W. Foundations of cost-effectiveness analysis for health and medical practices. *N Eng J Med* 1977;**296**:716-21.

43. Richardson G, Manca A. Calculation of Quality Adjusted Life Years (QALYs) in the published literature: a review of methodology and transparency. *Health Econ* 2004;13:1203-10.
44. Manca A, Hawkins N, Sculpher M. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health Econ* 2005;14:487-96.
45. Pliskin J, Shephard DS, Weinstein M. Utility functions for life years and health status. *Oper Res* 1980;28:206-24.
46. Williams A. Integration equity: an exploration of the 'fair innings' argument. *Health Econ* 1997;6:117-32.
47. Tsuchiya A, Williams A. A "fair innings" between the sexes: are men being treated inequitably? *Soc Sci Med* 2004;60:277-86.
48. Wailoo A, Anand P. The nature of procedural preferences for health-care rationing decisions. *Soc Sci Med* 2005;60:223-6.
49. Anand P, Wailoo A. Utilities versus rights for publicly provided goods: arguments and evidence from health care rationing. *Economica* 2000;67:543-77.
50. Freemantle N, Mason J. Not playing with a full DEC: why development and evaluation committee methods for appraising new drugs may be inadequate. *BMJ* 1999;318:1480-2.
51. McGregor M. Cost-utility analysis: use QALYs only with great caution. *CMAJ* 2003;168:433-4.
52. Harris J. It's not NICE to discriminate. *J Med Ethics* 2005;31:373-5.
53. Stevens A. Playing with a full DEC: quick, clean and fit for purpose [letter]. *BMJ* 1999;318:1480-2.
54. Nozick R. *Anarchy, State and Utopia*. Oxford: Basil Blackwell; 1974.
55. Richardson J, Abel Olsen J, Hawthorne G, Mortimer D, Smith R. *The measurement and valuation of quality of life in economic evaluation (CHPE Working Paper 97)*. Melbourne: Centre for Health Program Evaluation; 1998.
56. Sheldon T. It's just evaluation for decision making. In: Smith P, Ginnelly L, Sculpher M, editors. *Health Policy and Economics: Opportunities and Challenges*. Oxford: Oxford University Press; 2005. p. 33-36.
57. Stant D, Buskens E, Jenner J, Wiersma D, TenVergert E. Cost-effectiveness analysis in severe mental illness: outcome measures selection. *J Ment Health Policy Econ* 2007;10:101-8.
58. Sen A. *On ethics and economics*. Oxford: Basil Blackwell; 1987.
59. Phelps CE, Mushlin AI. On the (near) equivalence of cost-effectiveness and cost-benefit analyses. *Int J Technol Assess Health Care* 1991;7:12-21.
60. Dolan P, Edlin R. Is it really possible to build a bridge between cost-benefit analysis and cost-effectiveness analysis? *J Health Econ* 2002;21:827-43.
61. Culyer AJ, McCabe C, Briggs A, Claxton K, Buxton M, Akehurst R, et al. Searching for a threshold, not setting one: the role of the National Institute of Health and Clinical Excellence. *J Health Serv Res Policy* 2007;12:56-8.
62. Sculpher MJ, Claxton K, Drummond M, McCabe C. Whither trial-based economic evaluation for health care decision making? *Health Econ* 2006;15:677-87.
63. Mark D, Hlatky M, Califf R, Naylor C, Lee K, Armstrong P, et al. Cost effectiveness of thrombolytic therapy with tissue plasminogen activator as compared with streptokinase for acute myocardial infarction. *N Engl J Med* 1995;332:1418-24.
64. Nadiga D, Hensley P, Uhlenhuth E. Review of the long term effectiveness of cognitive behavioural therapy compared to medications in panic disorder. *Depress Anxiety* 2003;17:58-64.

65. Sculpher M. Evaluating the cost-effectiveness of interventions designed to increase the utilization of evidence-based guidelines. *Fam Pract* 2000;17:S26-31.
66. Kennedy A, Reeves D, Bower P, Lee V, Middleton E, Richardson G, et al. The effectiveness and cost effectiveness of a national lay led self care support programme for patients with long-term conditions: a pragmatic randomised controlled trial. *J Epidemiol Community Health* 2007;61:254-61.
67. Bojke L, Claxton K, Palmer S, Sculpher M. *Defining and characterising structural uncertainty in decision analytic models (CHE Research Paper 9)*. York: Centre for Health Economics; 2006.
68. Briggs A. Handling uncertainty in cost-effectiveness models. *Pharmacoeconomics* 2000;17:479-500.
69. McCabe C, Dixon S. Testing the validity of cost-effectiveness models. *Pharmacoeconomics* 2000;17:501-13.
70. Draper D. Assessment and propagation of model uncertainty. *J R Stat Soc* 1995;57:45-97.
71. Lorig K, Holman H. Evidence suggesting that health education for self-management in patients with chronic arthritis has sustained health benefits while reducing health care costs. *Arthritis Rheum* 1993;439-446.
72. Lorig K. Self-management education: more than a nice extra.[comment]. *Med Care* 2003;41:699-701.
73. Lorig K, Gonzalez VM, Ritter P. Community-based Spanish language arthritis education program: a randomized trial. *Med Care* 1999;37:957-63.
74. Lorig K, Ritter PL, Plant K. A disease-specific self-help program compared with a generalized chronic disease self-help program for arthritis patients. *Arthritis Rheum* 2005;53:950-7.
75. Lorig KR, Ritter P, Stewart AL, Sobel DS, Brown BW, Bandura A, et al. Chronic disease self-management program: 2-year health status and health care utilization outcomes. *Med Care* 2001;39:1217-23.
76. Lorig KR, Ritter PL, Gonzalez VM. Hispanic chronic disease self-management: a randomized community-based outcome trial. *Nurs Res* 2003;52:361-9.
77. Lorig KR, Ritter PL, Laurent DD, Fries JF. Long-term randomized controlled trials of tailored-print and small-group arthritis self-management interventions. *Med Care* 2004;42:346-54.
78. Lorig KR, Ritter PL, Laurent DD, Plant K. Internet-based chronic disease self-management: a randomized trial. *Med Care* 2006;44:964-71.
79. Lorig KR, Sobel DS, Ritter PL, Laurent D, Hobbs M. Effect of a self-management program on patients with chronic disease. *Eff Clin Pract* 2001;4:256-62.
80. Lorig KR, Sobel DS, Stewart AL, Brown BW, Bandura A, Ritter P, et al. Evidence suggesting that a chronic disease self-management program can improve health status while reducing hospitalization: a randomized trial. *Med Care* 1999;37:5-14.
81. Fu D, Fu H, McGowan P, Shen YE, Zhu L, Yang H, et al. Implementation and quantitative evaluation of chronic disease self-management programme in Shanghai, China: randomized controlled trial. *Bull World Health Organ* 2003;81:174-82.
82. Siu AM, Chan CC, Poon PK, Chui DY, Chan SC. Evaluation of the chronic disease self-management program in a Chinese population. *Patient Educ Couns* 2007;65:42-50.

83. Griffiths C, Motlib J, Azad A, Ramsay J, Eldridge S, Feder G, et al. Randomised controlled trial of a lay-led self-management programme for Bangladeshi patients with chronic disease. *Br J Gen Pract* 2005;55:831-7.
84. Ades A, Sculpher M, Sutton A, Abrams K, Cooper N, Welton N, et al. Bayesian methods for evidence synthesis in cost-effectiveness analysis. *Pharmacoeconomics* 2006;24:2-29.
85. Annis C. *Frequentists and Bayesians. What is "probability?"* 2006 [cited 2007 14 Nov]. Available from: http://www.statisticalengineering.com/frequentists_and_bayesians.htm
86. Sheingold S. Can Bayesian methods make data and analyses more relevant to decision makers? A perspective from Medicare. *Int J Technol Assess Health Care* 2001;17:114-22.
87. Ashby D, Smith A. Evidence-based medicine as Bayesian decision-making. *Stat Med* 2000;19:3291-302.
88. Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health Econ* 2001;10:779-87.
89. Claxton K, Sculpher M, McCabe C, Briggs A, Akehurst R, Buxton M, et al. Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra. *Health Econ* 2005;14:339-47.
90. van Hout BA, Al MJ, Gordon GS, Rutten FF. Costs, effects and C/E-ratios alongside a clinical trial. *Health Econ* 1994;3:309-19.
91. Claxton K, Cohen J, Neumann P. When is evidence sufficient? A framework for making use of all available information in medical decision making and for deciding whether more is needed. *Health Aff* 2005;24:93-101.
92. Claxton K, Sculpher M, Drummond M. A rational framework for decision making by the National Institute for Clinical Excellence. *Lancet* 2002;360:711-5.
93. Altman D, Bland M. Statistics notes: Absence of evidence is not evidence of absence. *BMJ* 1995;311:485.
94. Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *J Health Econ* 1999;18:341-64.
95. Austin P, Naylor D, Tu J. A comparison of a Bayesian vs. a frequentist method for profiling hospital performance. *J Eval Clin Pract* 2001;7:35-45.
96. Cooper N, Sutton A, Mugford M, Abrams K. Use of Bayesian Markov Chain Monte Carlo methods to model cost of illness data. *Med Decis Making* 2003;23:38-53.
97. Spiegelhalter D, Thomas A, Best N, Lunn D. *WinBUGS 1.4 User Manual*. 2003 [cited 2007 14 Nov]. Available from: <http://www.mrc-bsu.cam.ac.uk/bugs>
98. Kynn M. *Missing data in WinBUGS*. 2006 [cited 2007 4 Apr]. Available from: http://www.cas.lancs.ac.uk/short_courses/notes/missingdata/MissingBayesian.article.pdf
99. Drummond M, O'Brien B, Stoddart G. *Methods for the economic evaluation of health care programmes*. 2nd ed. Oxford: Oxford University Press; 1997.
100. Gravelle H, Richardson G, Weatherly H, Ritchie G. *An assessment of the quality of economic evaluations of self management*. York: Centre for Health Economics, University of York; 2003.
101. Weingarten SR, Henning JM, Badamgarav E, Knight K, Hasselblad V, Gano A, et al. Interventions used in disease management programmes for patients with chronic illness-which ones work? Meta-analysis of published reports. *BMJ* 2002;325:925.

102. Vallance A. Something out of nothing: the placebo effect. *Advances in Psychiatric Treatment* 2006;12:287-296.
103. Rogers A, Bower P, Gravelle H. *Evaluating the expert patient programme*. Manchester: National Primary Care Research and Development Centre, University of Manchester; 2002.
104. Riemsma RP, Taal E, Kirwan JR, Rasker JJ. Patient education programmes for adults with rheumatoid arthritis: benefits are small and short lived. *BMJ* 2002;325:558-9.
105. Kauppinen R, Vilkkä V, Sintonen H, Klaukka T, Tukiainen H. Long-term economic evaluation of intensive patient education during the first treatment year in newly diagnosed adult asthma. *Respir Med* 2001;95:56-63.
106. Ament A, Evers S, Goossens M, de Vet H., van Tulder M. Criteria list for conducting systematic reviews based on economic evaluation studies. In: Donaldson C, Mugford M, Vale L, editors. *Evidence-based Health Economics*. London: BMJ Books; 2002. p. 99-113.
107. Chiou CF, Hay JW, Wallace JF, Bloom BS, Neumann PJ, Sullivan SD, et al. Development and validation of a grading system for the quality of cost-effectiveness studies. *Med Care* 2003;41:32-44.
108. Centre for Reviews and Dissemination. Improving access to cost-effectiveness information for health care decision-making. The NHS Economic Evaluation Database. Report 6. 2nd ed. York: NHS Centre for Reviews and Dissemination, University of York; 2001.
109. Forbes C, Wilby J, Richardson G, Sculpher M, Mather L, Riemsma R. A systematic review and economic evaluation of pegylated liposomal doxorubicin hydrochloride for ovarian cancer. *Health Technol Assess* 2002;6:1-119.
110. Mauskopf J, Paul J, Grant DM, Stergachis A. The role of cost-consequence analysis in healthcare decision-making. *Pharmacoeconomics* 1998;13:277-88.
111. Albisser AM, Harris RI, Albisser JB, Sperlich M. The impact of initiatives in education, self-management training, and computer-assisted self-care on outcomes in diabetes disease management. *Diabetes Technology & Therapeutics* 2001;3:571-579.
112. Berg GD, Wadhwa S. Diabetes disease management in a community-based setting. *Managed Care* 2002;11:42-50.
113. Cline CMJ, Israelsson BYA, Willenheimer RB, Broms K, Erhardt LR. Cost effective management programme for heart failure reduces hospitalisation. *Heart* 1998;80:442-446.
114. Lafata JE, Martin SA, Kaatz S, Ward RE. Anticoagulation clinics and patient self-testing for patients on chronic warfarin therapy: A cost-effectiveness analysis. *Journal of Thrombosis & Thrombolysis* 2000;9 Suppl 1:S13-9.
115. Engh CA, Jr., Culpepper WJ, 2nd, Charette PA, Brown R. Patient self-testing of prothrombin time after hip arthroplasty. *Journal of the Southern Orthopaedic Association* 2001;10:140-6.
116. Fitzmaurice DA, Murray ET, Gee KM, Allan TF, Hobbs FD. A randomised controlled trial of patient self management of oral anticoagulation treatment compared with primary care management. *J Clin Pathol* 2002;55:845-9.
117. Gallefoss F, Bakke PS. Cost-effectiveness of self-management in asthmatics: a 1-yr follow-up randomized, controlled trial. *European Respiratory Journal* 2001;17:206-213.

118. Gallefoss F, Bakke PS. Cost-benefit and cost-effectiveness analysis of self-management in patients with COPID: a 1-year follow-up randomized, controlled trial. *Respiratory Medicine* 2002;**96**:424-431.
119. Ghosh C, Ravindran P, Joshi M, C SS. Reductions in hospital use from self management training for chronic asthmatics. *Social Science and Medicine* 1998;**46**:1087-1093.
120. Glasgow RE, La Chance PA, Toobert DJ, Brown J, Hampson SE, Riddle MC. Long term effects and costs of brief behavioural dietary Intervention for patients with diabetes delivered from the medical office. *Patient Education & Counseling* 1997;**32**:175-184.
121. Gray A, Raikou M, McGuire A, Fenn P, Stevens R, Cull C, et al. Cost effectiveness of an intensive blood glucose control policy in patients with type 2 diabetes: economic analysis alongside randomised controlled trial (UKPDS 41). *BMJ* 2000;**320**:1373-8.
122. Groessl EJ, Cronan TA. A cost analysis of self-management programs for people with chronic illness. *American Journal of Community Psychology* 2000;**28**:455-480.
123. Humphreys K, Moos R. Can encouraging substance abuse patients to participate in self-help groups reduce demand for health care? A quasi-experimental study. *Alcoholism: Clinical & Experimental Research* 2001;**25**:711-716.
124. Jacobsen PB, Meade CD, Stein KD, Chirikos TN, Small BJ, Ruckdeschel JC. Efficacy and costs of two forms of stress management training for cancer patients undergoing chemotherapy. *Journal of Clinical Oncology* 2002;**20**:2851-2862.
125. Kauppinen R, Sintonen H, Tukiainen H. One-year economic evaluation of intensive vs conventional patient education and supervision for self-management of new asthmatic patients. *Respiratory Medicine* 1998;**92**:300-307.
126. Kauppinen R, Sintonen H, Vilkkä V, Tukiainen H. Long-term (3-year) economic evaluation of intensive patient education for self-management during the first year in new asthmatics. *Respiratory Medicine* 1999;**93**:283-289.
127. Kruger J, Helmick C G, Callahan L F, C HA. Cost-effectiveness of the arthritis self-help course. *Archives of Internal Medicine* 1998;**158**:1245-1249.
128. Lahdensuo A, Haahtela T, Herrala J, Kava T, Kiviranta K, Kiuusisto P, et al. Randomised comparison of cost effectiveness of guided self management and traditional treatment of asthma in Finland. *British Medical Journal* 1998;**316**:p 1138-9.
129. Lord J, Victor C, Littlejohns P, Ross FM, Axford JS. Economic evaluation of a primary care-based education programme for patients with osteoarthritis of the knee. *Health Technol Assess* 1999;**3**:1-55.
130. Mattson-Prince J. A rational approach to long-term care: comparing the independent living model with agency-based care for persons with high spinal cord injuries. *Spinal Cord* 1997;**35**:326-31.
131. Mazzuca SA, Brandt KD, Katz BP, Hanna MP, Melfi CA. Reduced utilization and cost of primary care clinic visits resulting from self-care education for patients with osteoarthritis of the knee. *Arthritis Rheum* 1999;**42**:1267-73.
132. Meier JL, Swislocki ALM, Lopez JR, Noth RH, Bartlebaugh P, Siegel D. Reduction in self-monitoring of blood glucose in persons with type 2 diabetes results in cost savings and no change in glycemic control. *American Journal of Managed Care* 2002;**8**:557-565.

133. Neri M, Migliori GB, Spanevello A, Berra D, Nicolin E, Landoni CV, et al. Economic analysis of two structured treatment and teaching programs on asthma. *Allergy* 1996;51:313-9.
134. Robinson A, Thompson DG, Wilkin D, Roberts C. Guided self-management and patient-directed follow-up of ulcerative colitis. *Lancet* 2001;358:967-81.
135. Rubin N, Foxman B. The cost-effectiveness of placing urinary tract infection treatment Over the counter. *Journal of Clinical Epidemiology* 1996;49:1315-1321.
136. Salkeld G, Phongsavan P, Oldenburg B, Johannesson M, Convery P, Graham-Clarke P, et al. The cost-effectiveness of a cardiovascular risk reduction program in general practice. *health Policy* 1997;41:105-119.
137. Schermer TR, Thoonen BP, van den Boom G, Akkermans RP, Grol RP, Folgering HT, et al. Randomized controlled economic evaluation of asthma self-management in primary health care. *American Journal of Respiratory & Critical Care Medicine* 2002;166:1062-1072.
138. Sinclair HK, Silcock J, Bond CM, Lennox AS, Winfield AJ. The cost-effectiveness of intensive pharmaceutical intervention in assisting people to stop smoking. *Int J Pharm Pract* 1999;7:107-12.
139. Starostina EG, Antsiferov M, Galstyan GR, Trautner C, Jorgens V, Bott U, et al. Effectiveness and Cost-Benefit Analysis of Intensive Treatment and Teaching Programmes for Type 1 (Insulin- Dependent) Diabetes Mellitus in Moscow - Blood Glucose Versus Urine Glucose Self-Monitoring. *Diabetologia* 1994:170-176.
140. Taborski U, Wittstamm FJ, Bernardo A. Cost-effectiveness of self-managed anticoagulant therapy in Germany. *Seminars in Thrombosis & Hemostasis* 1999;25:103-107.
141. Tschopp J, Frey J, Pernet R, Burrus C, Jordan B, Morin A, et al. Bronchial asthma and self-management education: implementation of Guidelines by an interdisciplinary programme in a health network: study of respiratory education group (REG). *Swiss Medical Weekly* 2002;132:92-97.
142. Volsko TA. A pediatric asthma clinic pilot program reduces emergency department visits, hospitalizations, and cost of care. *Respiratory Care* 1998;43:107-113.
143. Von Korff MBWCDDRA. Effects of Practice Style in Managing Back Pain. *Annals of Internal Medicine*. 1994;121:187-195.
144. Watson MC, Bond CM, Grimshaw JM, Mollison J, Ludbrook A, Walker AE. Educational strategies to promote evidence-based community pharmacy practice: a cluster randomized controlled trial (RCT). *Fam Pract* 2002;19:529-36.
145. Castro M, Zimmermann NA, Crocker S, Bradley J, Leven C, Schechtmann KB. Asthma intervention program prevents readmissions in high healthcare costs. *Am J Respir Crit Care Med* 2003;168:1095-1099.
146. Dunagan WC, Littenberg B, Ewald GA, Jones CA, Emery VB, Waterman BM, et al. Randomized trial of a nurse-administered, telephone-based disease management program for patients with heart failure. *Journal of Cardiac Failure* 2005;11:358-365.
147. Gallefoss F. The effects of patient education in COPD in a 1-year follow-up randomised, controlled trial. *Patient Education & Counseling* 2004;52:259-66.
148. Jowett S, Bryan S, Murray E, McCahon D, Raftery J, Richard Hobbs FD, et al. Patient self-management of anticoagulation therapy: a trial-based cost-effectiveness analysis. *Br J Haematol* 2006;134:632-9.

149. Keers JC, Groen H, Sluiter WJ, Bouma J, Links TP. Cost and benefits of a multidisciplinary intensive diabetes education programme. *Journal of Evaluation in Clinical Practice* 2005;11:293-303.
150. Lindberg M, Ahlner J, Ekstrom T, Jonsson D, Moller M. Asthma nurse practice improves outcomes and reduces costs in primary health care. *Scandinavian Journal of Caring Sciences* 2002;16:73-78.
151. Monninkhof E, van der Valk P, Schermer T, van der Palen J, van Herwaarden C, Zielhuis G. Economic evaluation of a comprehensive self-management programme in patients with moderate to severe chronic obstructive pulmonary disease. *Chronic Respiratory Disease* 2004;1:7-16.
152. Mortimer D, Kelly J. Economic evaluation of the good life club intervention for diabetes self-management. *Australian Journal of Primary Health* 2006;12:91-100.
153. Richards A, Barkham M, Cahill J, Richards D, Williams C, Heywood P. PHASE: a randomised, controlled trial of supervised self-help cognitive behavioural therapy in primary care. *Br J Gen Pract* 2003;53:764-70.
154. Richardson G, Sculpher M, Kennedy A, Nelson E, Reeves D, Roberts C, et al. Is self-care a cost-effective use of resources? Evidence from a randomized trial in inflammatory bowel disease. *J Health Serv Res Policy* 2006;11:225-30.
155. Robinson A, Lee V, Kennedy A, Middleton L, Rogers A, Thompson DG, et al. A randomised controlled trial of self-help interventions in patients with a primary care diagnosis of irritable bowel syndrome. *Gut* 2006;55:643-8.
156. Strong LL, Von Korff M, Saunders K, Moore JE. Cost-effectiveness of two self-care interventions to reduce disability associated with back pain. *Spine* 2006;31:1639-45.
157. Tagaya E, Tamaoki J, Nagai A, Murasugi H, Igi H. The role of a self-management program in the control of mild to moderate asthma: A randomized controlled study. *Allergology International* 2005;54:pp 527-531.
158. Waxman R, Woodburn H, Powell M, Woodburn J, Blackburn S, Helliwell P. FOOTSTEP: a randomized controlled trial investigating the clinical and cost effectiveness of a patient self-management program for basic foot care in the elderly. *J Clin Epidemiol* 2003;56:1092-9.
159. Manca A, Palmer S. Handling missing data in patient-level cost-effectiveness analysis alongside randomised clinical trials. *Appl Health Econ Health Policy* 2005;4:65-75.
160. Barlow J, Wright C, Sheasby J, Turner A, Hainsworth J. Self-management approaches for people with chronic conditions: a review. *Patient Educ Couns* 2002;48:177-87.
161. Bower P, Richards D, Lovell K. The clinical and cost-effectiveness of self-help treatments for anxiety and depressive disorders in primary care: a systematic review. *Br J Gen Pract* 2001;51:838-45.
162. Richardson G, Kennedy A, Reeves D, Bower P, Lee V, Middleton E, et al. Cost-effectiveness analysis of a national lay led self care support programme for patients with long-term conditions. *J Epidemiol Community Health* 2007;(in press).
163. Culyer AJ. NICE's use of cost effectiveness as an exemplar of a deliberative process. *Health Econ Pol Law* 2006;1:299-318.
164. Department of Health. *Our Health, our care, our say: making it happen*. London: Department of Health; 2006.
165. Delaney B. Economic analysis and complex randomized controlled trials. *Family Practice* 2007;24:293-294.

166. Cooper J, Thompson J. *Stepping stones to success: an implementation, training and support framework for lay led self-management programmes*. London: NHS; 2005.
167. Medical Research Council. *A framework for development and evaluation of RCTs for complex interventions to improve health*. London: Medical Research Council; 2000.
168. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. *Psychol Rev* 1977;84:191-215.
169. Netten A, Curtis L. Unit Costs of Health and Social Care 2002. Canterbury: Personal Social Services Research Unit, University of Kent; 2002.
170. Curtis L, Netten A. Unit costs of health and social care 2004. Canterbury: Personal Social Services Unit, University of Kent; 2004.
171. British Medical Association. *British National Formulary 48*. London: British Medical Association; 2004.
172. Department of Health. Personal Communication. 2005.
173. Kind P. The EuroQoL instrument: an index of health-related quality of life. In: Spilker B, editor. *Quality of life and pharmacoeconomics in clinical trials*. Philadelphia: Lippincott-Raven; 1996. p. 191-201.
174. Dolan P, Gudex C, Kind P, Williams A. A social tariff for EuroQol: results from a UK general population survey (CHE discussion paper no.138). York: Centre for Health Economics; 1995.
175. Matthews J, Altman D, Campbell M, Royston P. Analysis of serial measurements in medical research. *BMJ* 1990;300:230-5.
176. Rubin DB, Stern HS, Vehovar V. Handling "don't know" survey responses: the case of the Slovenian plebiscite. *J Am Stat Assoc* 1995;90:822-8.
177. Stinnett AA, Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. *Med Decis Making* 1998;18:S68-80.
178. Briggs A, Wonderling D, Mooney C. Pulling cost-effectiveness analysis up by its bootstraps: a non-parametric approach to confidence interval estimation. *Health Econ* 1997;6:327-40.
179. Stata. Statistical Data Analysis. www.stata.com 2002;volume 8.0.
180. National Centre for Social Research and University College London. *Health Survey for England*. Colchester: National Centre for Social Research; 2005.
181. Simmonds M, Higgins J, Stewart L, Tierney J, Clarke M, Thompson S. Meta-analysis of individual patient data from randomized trials: a review of methods used in practice. *Clin Trials* 2005;2:209-17.
182. Berlin J, Santanna J, Schmid C, Szczech L, Feldman H. Individual patient versus group-level data meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head. *Stat Med* 2002;21:371-87.
183. Sutton A, Duval S, Tweedie R, Abrams K, Jones D. Empirical assessment of effect of publication bias on meta-analyses. *BMJ* 2000;320.
184. Thompson S, Higgins J. Treating individuals 4: can meta-analysis help target interventions at individuals most likely to benefit? *Lancet* 2005;365:341-6.
185. Lambert P, Sutton A, Abrams K, Jones D. A comparison of summary patientlevel covariates in meta-regression with individual patient data meta-analysis. *J Clin Epidemiol* 2002;55:86-94.
186. Willan A, Briggs A, Hoch J. Regression methods for covariate adjustment and subgroup analysis for non-censored cost-effectiveness data. *Health Econ* 2004;13:461-75.

187. Sculpher M, Claxton K. Establishing the cost-effectiveness of new pharmaceuticals under conditions of uncertainty. When is there sufficient evidence? *Value Health* 2005;8:433-46.
188. Ades A, Claxton K, Sculpher M. Evidence synthesis, parameter correlation and probabilistic sensitivity analysis. *Health Econ* 2006;15:373-81.
189. Hawkins N, Richardson G, Sutton A, N C. A comparison of direct, indirect and joint analyses of cost-effectiveness. *HESG paper, Brunel* 2007 2007.
190. Higgins J, Whitehead A, Turner R, Omar R, Thompson S. Metaanalysis of continuous outcome data from individual patients. *Stat Med* 2001;20:2219-41.
191. Stewart L, Tierney J. To IPD or not to IPD? *Eval Health Prof* 2002;25:76-97.
192. Advanced Bladder Cancer Meta Analysis Collaboration. Adjuvant chemotherapy in invasive bladder cancer: A systematic review and meta-analysis of individual patient data. *Eur Urol* 2005;48:189-201.
193. Non-small Cell Lung Cancer Collaborative Group. Chemotherapy in non-small cell lung cancer: a meta-analysis using updated data on individual patients from 52 randomised trials. *BMJ* 1995;311:899-909.
194. Sutton A, Kendrick D, Coupland C. Meta-analysis of individual and aggregate level data. *Stat Med* 2007;(in press).
195. Bandolier. *Surrogate endpoints- glossary definition*. 2004 [cited 2007 14 Nov]. Available from: <http://www.ir2.ox.ac.uk/bandolier/booth/glossary/surrog.html>
196. Fleming T, DeMets D. Surrogate endpoints in clinical trials: are we being misled? *Ann Intern Med* 1996;125:605-13.
197. Ellenberg S. Surrogate markers in AIDS and cancer trials- discussion. *Stat Med* 1994;13:1437-40.
198. Baker S, Kramer B. A perfect correlate does not a surrogate make. *BMC Med Res Methodol* 2003;3:1-5.
199. Weir C, Walley R. Statistical evaluation of biomarkers as surrogate endpoints: a literature review. *Stat Med* 2006;25:183-203.
200. De Gruttola V, Fleming T, Lin D, Coombs R. Perspective: validating surrogate markers—are we being naive? *J Infect Dis* 1997;175:237-46.
201. Lin D, Fleming D, De Gruttola V. Estimating the proportion of treatment effect explained by a surrogate marker. *Stat Med* 1997;16:1515-27.
202. Spiegelhalter D, Best N. Bayesian approaches to multiple sources of evidence and uncertainty in complex cost-effectiveness modelling. *Stat Med* 2003;22:3687-709.
203. Grieve R, Nixon R, Thompson S, Cairns J. Multilevel models for estimating incremental net benefit in multinational studies. *Health Econ* 2007;(in press).
204. Manca A, Willan A. Lost in translation. *Pharmacoeconomics* 2006;24:1-19.
205. Willke R, Glick H, Polsky D, Schulman K. Estimating the country specific cost-effectiveness from multinational clinical trials. *Health Econ* 1998;7:481-93.
206. Zwanziger J, Anderson G, Haber S, Thorpe K, Newhouse J. Comparison of hospital costs in California, New York, and Canada. *Health Aff* 1993;12:130-9.
207. Rouleau J, Moye L, Pfeffer M, Arnold J, Bernstein V, Cuddy T, et al. A comparison of management patterns after myocardial infarction in Canada and the United States. *N Engl J Med* 1993;328:779-84.
208. Mark D, Naylor C, Hlatky M. Use of medical resources and quality of life after acute myocardial infarction in Canada and the United States. *N Engl J Med* 1994;331:1130-5.
209. Ades A, Lu G, Claxton K. Expected value of sample information calculations in medical decision modeling. *Med Decis Making* 2004;24:207-27.

210. McAlister F. Commentary: Relative treatment effects are consistent across the spectrum of underlying risks ... usually. *International Journal of Epidemiology* 2002;31:76-77.
211. Chatfield C. Model uncertainty, data mining and statistical inference. *J R Stat Soc* 1995;158:419-66.
212. Snowling S, Kramer J. Evaluating modelling uncertainty for model selection. *Ecol Modell* 2001;138:17-30.
213. Solomon DH, Warsi A, Brown-Stevenson T, Farrell M, Gauthier S, Mikels D, et al. Does self-management education benefit all populations with arthritis? A randomized controlled trial in a primary care physician network. *J Rheumatol* 2002;29:362-8.
214. Yip YB, Sit JW, Fung KK, Wong DY, Chong SY, Chung LH, et al. Impact of an arthritis self-management programme with an added exercise component for osteoarthritic knee sufferers on improving pain, functional outcomes, and use of health care services: an experimental study. *Patient Educ Couns* 2007;65:113-21.
215. Barlow JH, Turner AP, Wright CC. A randomized controlled study of the arthritis self-management programme in the UK. *Health Educ Res* 2000;15:665-80.
216. Buszewicz M, Rait G, Griffin M, Nazareth I, Patel A, Atkinson A, et al. Self management of arthritis in primary care: randomised controlled trial. *BMJ* 2006;333:879.
217. LeFort S. *Cost effectiveness of the chronic pain self-management program [abstract]*. In: RCN International Nursing Research Conference; 2004; University of Cambridge, England; 2004. p. 54 abst 77.
218. Murphy MC, Berlowitz DJ, Saunders JE, Campbell M, Jackson B. A randomised trial to compare pulmonary rehabilitation (PRP) & the Stanford model chronic disease self management program (CDSMP) in COPD with 12 month follow-up [abstract]. *Respirology* 2005;10:A61.
219. Haas M, Group E, Muench J, Kraemer D, Brummel-Smith K, Sharma R, et al. Chronic disease self-management program for low back pain in the elderly. *J Manipulative Physiol Ther* 2005;28:228-37.
220. Swerissen H, Belfrage J, Weeks A, Jordan L, Walker C, Furler J, et al. A randomised control trial of a self-management program for people with a chronic illness from Vietnamese, Chinese, Italian and Greek backgrounds. *Patient Educ Couns* 2006;64:360-8.
221. Ryan M. Discrete choice experiments in health care. *BMJ* 2004;328:360-1.
222. Richardson G, Bojke C, Kennedy A, Reeves D, Bower P, Lee V, et al. What outcomes are important to patients with long term conditions? A discrete choice experiment. *Value in Health* 2007;(in press).
223. Richardson G. Self care interventions: which outcomes are important to consumers? In: 13th Cochrane Colloquium; 2005 22-26 October; Melbourne; 2005.
224. Richardson G, Bojke C, Bower P, Gardner C, Gately C, Lee V, et al. How do we value psychological outcomes in systematic reviews? In: 14th Cochrane Colloquium; 2006 23-26 October; Dublin; 2006.
225. Richardson G, Bojke, C. Incorporating "psychological" outcomes into cost-effectiveness analysis. *HESG paper, York 2006* 2006.
226. Department of Health. *The expert patient: a new approach to chronic disease management for the 21st century*. London: Department of Health; 2001.
227. Lancaster K. A new approach to consumer theory. *J Polit Econ* 1966;74:132-157.

228. Bergmann A, Hanley N, Wright R. Valuing the attributes of renewable energy investments. In: European Association of Environmental and Resource Economists conference; 2004; Budapest; 2004.
229. Manski C. The structure of random utility models. *Theory Decis* 1977;8:229-54.
230. Gyrd-Hansen D. Willingness to pay for a QALY. *Health Econ* 2003;12:1049-60.
231. Doward LC, Ratcliffe J, McKenna SP, Meads DM, Langley PC. A comparison of conjoint analysis and time trade off elicited utilities for quality of life states associated with the treatment of genital herpes. *Value Health* 2002;5:574.
232. Bala MV, Mauskopf J. Estimating the Bayesian loss function: a conjoint analysis approach. *Int J Technology Assess Health Care* 2001;17:27-37.
233. Ryan M, Hughes J. Using conjoint analysis to assess women's preferences for miscarriage management. *Health Econ* 1997;6:261-73.
234. Sculpher M, Bryan S, Fry P, de Winter P, Payne H, Emberton M. Patients' preferences for the management of non-metastatic prostate cancer: discrete choice experiment. *BMJ* 2004;328:382.
235. Fraenkel L, Bogardus Jr ST, Wittink DR. Risk-attitude and patient treatment preferences. *Lupus* 2003;12:370-6.
236. Gerard K, Lattimer V. Preferences of patients for emergency services available during usual GP surgery hours: a discrete choice experiment. *Fam Pract* 2005;22:28-36.
237. Gyrd-Hansen D. Cost-benefit analysis of mammography screening in Denmark based on discrete ranking data. *Int J Technol Assess Health Care* 2000;16:811-21.
238. Gyrd-Hansen D, Sogaard J. Analysing public preferences for cancer screening programmes. *Health Econ* 2001;10:617-34.
239. Hall J, Kenny P, King M, Louviere J, Viney R, Yeoh A. Using stated preference discrete choice modelling to evaluate the introduction of varicella vaccination. *Health Econ* 2002;11:457-65.
240. Hill AR, Aspinall P, Dhillon B, Ambrecht A, Lumsden C, Nelson P, et al. Utilities for quality of life from time trade-off and conjoint analysis in age related macular degeneration. *Invest Ophthalmol Vis Sci* 2004;45:3080.
241. Hundley V, Ryan M, Graham W. Assessing women's preferences for intrapartum care. *Birth* 2001;28:254-63.
242. Johnson FR, Banzhaf MR, Desvousges WH. Willingness to pay for improved respiratory and cardiovascular health: A multiple-format, stated-preference approach. *Health Econ* 2000;9:295-317.
243. Mark TL, Swait J. Using stated preference and revealed preference modeling to evaluate prescribing decisions. *Health Econ* 2004;13:563-73.
244. McKenzie L, Cairns J, Osman L. Symptom-based outcome measures for asthma: the use of discrete choice methods to assess patient preferences. *Health Policy* 2001;57:193-204.
245. Payne K, Elliott R. Using discrete choice experiments to value preferences for pharmacy services. *Int J Pharm Pract* 2005;13:9-20.
246. Petrou S, Henderson J. Preference-based approaches to measuring the benefits of perinatal care. *Birth* 2003;30:217-26.
247. Bandura A, Adams NE, Beyer J. Cognitive processes mediating behavioral change. *J Pers Soc Psychol* 1977;35:125-39.

248. Bandura A, Jeffery RW, Gajdos E. Generalizing change through participant modeling with self-directed mastery. *Behav Res Ther* 1975;13:141-52.
249. Viney R, Savage E, Louviere J. Empirical investigation of experimental design properties of discrete choice experiments in health care. *Health Econ* 2005;14:349-62.
250. Louviere JJ, Hensher DA, Swait JD. *Stated choice methods. Analysis and application*. Cambridge: Cambridge University Press; 2000.
251. Kocur G, Adler T, Hyman W, Aunet B. *Guide to forecasting travel demand with direct utility assessment*. Washington, DC: Urban Mass Transportation Administration, US Dept of Transportation; 1982. Report No.: UMTA-NH-11-0001-82-1.
252. Hahn G, Shapiro S. *A catalogue and computer program for the design and analysis of orthogonal symmetric and asymmetric fractional factorial experiments*. New York: General Electric Research and Development Centre; 1966. Report No.: 66-C-165.
253. Ryan M, Amaya-Amaya M. 'Threats' to and hopes for estimating benefits. *Health Econ* 2005;14:609-19.
254. Lloyd AJ. Threats to the estimation of benefit: Are preference elicitation methods accurate? *Health Econ* 2003;12:393-402.
255. Lancsar E, Savage E. Deriving welfare measures from discrete choice experiments: A response to Ryan and Santos Silva. *Health Econ* 2004;13:919-24.
256. Skjoldborg US, Gyrd-Hansen D. Conjoint analysis. The cost variable: An Achilles' heel? *Health Econ* 2003;12:479-91.
257. van der Pol M. *DCE with a contingent valuation follow-up*. In: 3rd Workshop Advancing the Methodology of Discrete Choice Experiments in Health Economics; 2005 6-7 July; University of Las Palmas de Gran Canaria: ECOMAS Research Unit; 2005.
258. Arana JE. *Emotions and decision rules in Discrete Choice Experiments for valuing health care programmes for the elderly*. In: 3rd Workshop Advancing the Methodology of Discrete Choice Experiments in Health Economics; 2005 6-7 July; University of Las Palmas de Gran Canaria: ECOMAS Research Unit; 2005.
259. Dawes R, Corrigan B. Linear models in decision making. *Psychol Bull* 1974;81:95-106.
260. Flynn T. *Best-worst scaling: estimating preferences for aspects of a dermatology consultation*. In: 3rd Workshop Advancing the Methodology of Discrete Choice Experiments in Health Economics; 2005 6-7 July; University of Las Palmas de Gran Canaria: ECOMAS Research Unit; 2005.
261. Ryan M, Gerard K. Using discrete choice experiments to value health care programmes: current practice and future research reflections. *Appl Health Econ Health Policy* 2003;2:55-64.
262. Ryan M, Farrar S. Using conjoint analysis to elicit preferences for health care. *BMJ* 2000;320:1530-1533.
263. Ryan M, Tinelli M, M O. *Deriving welfare estimates in discrete choice experiments with multiple choice options*. In: 3rd Workshop Advancing the Methodology of Discrete Choice Experiments in Health Economics; 2005 6-7 July; University of Las Palmas de Gran Canaria: ECOMAS Research Unit; 2005.
264. Hanley N, Ryan M, Wright R. Estimating the monetary value of health care: lessons from environmental economics. *Health Econ* 2003;12:3-16.
265. Lancsar E. *Deleting "irrational" responses from stated choice experiments: a case of investigating or imposing preferences?* In: 3rd Workshop Advancing the

- Methodology of Discrete Choice Experiments in Health Economics; 2005 6-7 July; University of Las Palmas de Gran Canaria: ECOMAS Research Unit; 2005.
266. Lancsar E, Savage E. Deriving welfare measures from discrete choice experiments: Inconsistency between current methods and random utility and welfare theory. *Health Econ* 2004;**13**:901-7.
267. Santos Silva JMC. Deriving welfare measures in discrete choice experiments: a comment to Lancsar and Savage. *Health Econ* 2004;**13**:913-8.
268. San Miguel F. *Testing the assumptions of completeness, stability and rationality of preferences in health economics using discrete choice experiments* [PhD thesis]. Aberdeen: University of Aberdeen; 2000.
269. San Miguel F, Ryan M, Amaya-Amaya M. 'Irrational' stated preferences: a quantitative and qualitative investigation. *Health Econ* 2005;**14**:307-22.
270. Norwood FB. Can Calibration Reconcile Stated and Observed Preferences? *Journal of Agricultural and Applied Economics*, April 2005.
271. Cookson R. Incorporating psycho-social considerations into health valuations: an experimental study. *J Health Econ* 2000;**19**:369-401.
272. Brazier J, Akehurst R, Brennan A, Dolan P, Claxton K, McCabe C, et al. Should patients have a greater role in valuing health States? *Appl Health Econ Health Policy* 2005;**4**:201-8.
273. Dolan P. Whose preferences count? *Med Decis Making* 1999;**19**:482-6.

Appendices

Appendix A. *Search strategy*

NHS EED 1995-16.4.03

Accessed via CAIRS T system

Searched 16.4.03

NHS EED 2003-31.08.05 (160 records)

Accessed via CAIRS T system

Searched 31.08.05

S self(w)care

S self(w)manag\$

S self(w) monitor\$

S self(w)help

S self(w)treat\$

S (self(w)administer\$ andnot (self (w)administer\$ (2w)questionnaire\$ or self(w)administer\$(2w)interview\$))

S self(w)medicat\$

S self(w)diagnos\$

S group\$(w)support\$

S peer(w)support\$

S expert(w)patient\$

S ((pharmacist\$ or pharmacy or pharmacies) (2w) support\$)

S ((pharmacist\$ or pharmacy or pharmacies)(2w) assist\$)

S ((pharmacist or pharmacy or pharmacies)(2w) (advice or advise\$))

S pharmaceutical(w)care

S s1 or s2 or s3 or s4 or s5 or s6 or s7 or s8 or s9 or s10 or s11 or s12 or s13 or s14 or s15

S@2005SEP1:2007APR3

S s16 & s17

NHS EED 01.09.05-03.04.07 (169 records/ 166 after de-duplication)

Accessed via CAIRS T system

Searched 03.04.07

S self(w)care

S self(w)manag\$

S self(w) monitor\$

S self(w)help

S self(w)treat\$

S (self(w)administer\$ andnot (self (w)administer\$ (2w)questionnaire\$ or self(w)administer\$(2w)interview\$))

S self(w)medicat\$

S self(w)diagnos\$

S group\$(w)support\$

S peer(w)support\$

S expert(w)patient\$

S ((pharmacist\$ or pharmacy or pharmacies) (2w) support\$)
S ((pharmacist\$ or pharmacy or pharmacies)(2w) assist\$)
S ((pharmacist or pharmacy or pharmacies)(2w) (advice or advise\$))
S pharmaceutical(w)care
S s1 or s2 or s3 or s4 or s5 or s6 or s7 or s8 or s9 or s10 or s11 or s12 or s13 or s14
or s15
S@2005SEP1:2007APR3
S s16 & s17

DARE 1995-16.4.03

Accessed via CAIRS T system
Searched 16.4.03

DARE 2003-31.08.05 (records)

Accessed via CAIRS T system
Searched 31.08.05

S self(w)care
S self(w)manag\$
S self(w) monitor\$
S self(w)help
S self(w)treat\$
S (self(w)administer\$ andnot (self (w)administer\$ (2w)questionnaire\$ or
self(w)administer\$(2w)interview\$))
S self(w)medicat\$
S self(w)diagnos\$
S group\$(w)support\$
S peer(w)support\$
S expert(w)patient\$
S ((pharmacist\$ or pharmacy or pharmacies) (2w) support\$)
S ((pharmacist\$ or pharmacy or pharmacies)(2w) assist\$)
S ((pharmacist or pharmacy or pharmacies)(2w) (advice or advise\$))
S pharmaceutical(w)care
S s1 or s2 or s3 or s4 or s5 or s6 or s7 or s8 or s9 or s10 or s11 or s12 or s13 or s14
or s15
S econom\$ or cost or costs or costly or costing or price or prices or pricing or
pharmacoeconomics
S expenditure\$ andnot energy
S value (1w) money
S budget\$
S s17 or s18 or s19 or s20
S s16 and s21

DARE 01.09.05-03.04.07 (30 records/ 25 after de-deuplication)

Accessed via CAIRS T system
Searched 03.04.07

S self(w)care

S self(w)manag\$
 S self(w) monitor\$
 S self(w)help
 S self(w)treat\$
 S (self(w)administer\$ andnot (self (w)administer\$ (2w)questionnaire\$ or
 self(w)administer\$(2w)interview\$))
 S self(w)medicat\$
 S self(w)diagnos\$
 S group\$(w)support\$
 S peer(w)support\$
 S expert(w)patient\$
 S ((pharmacist\$ or pharmacy or pharmacies) (2w) support\$)
 S ((pharmacist\$ or pharmacy or pharmacies)(2w) assist\$)
 S ((pharmacist or pharmacy or pharmacies)(2w) (advice or advise\$))
 S pharmaceutical(w)care
 S s1 or s2 or s3 or s4 or s5 or s6 or s7 or s8 or s9 or s10 or s11 or s12 or s13 or s14
 or s15
 S econom\$ or cost or costs or costly or costing or price or prices or pricing or
 pharmacoeconomics
 S expenditure\$ andnot energy
 S value (1w) money
 S budget\$
 S s17 or s18 or s19 or s20
 S s16 and s21
 S@2005SEP1:2007APR3
 S s22 & s23

Health Technology Assessment Database

Accessed via <http://agatha.york.ac.uk/welcome.htm>

Searched 17.4.03

self-care or self-help-groups/Subject Headings Exploded
 OR self(s)care or self(s)manag or self(s)monitor or self(s)help or self(s)treat or
 self(s)administer or self(s)medicat or self(s)diagnos or group(s)support or expert
 patient or pharmaceutical care/All fields
 OR ((pharmacist or pharmacy or pharmacies) (s) (support or assist or advice or
 advise))/All fields

*Health Technology Assessment Database 01.09.05-03.04.07 (17 records/15 after de-
duplication)*

Accessed via CAIRS B system

Searched 03.04.07

S self(w)care
 S self(w)manag\$
 S self(w) monitor\$
 S self(w)help
 S self(w)treat\$

S (self(w)administer\$ andnot (self (w)administer\$ (2w)questionnaire\$ or
 self(w)administer\$(2w)interview\$))
 S self(w)medicat\$
 S self(w)diagnos\$
 S group\$(w)support\$
 S peer(w)support\$
 S expert(w)patient\$
 S ((pharmacist\$ or pharmacy or pharmacies) (2w) support\$)
 S ((pharmacist\$ or pharmacy or pharmacies)(2w) assist\$)
 S ((pharmacist or pharmacy or pharmacies)(2w) (advice or advise\$))
 S pharmaceutical(w)care
 S s1 or s2 or s3 or s4 or s5 or s6 or s7 or s8 or s9 or s10 or s11 or s12 or s13 or s14
 or s15
 S@2005SEP1:2007APR3
 S s16 & s17

National Research Register Issue 1 2003

Accessed via CDROM
 Searched 22.4.03

National Research Register Issue 3 2005

Accessed via Internet
 Searched 31.08.05
 Not possible to limit by date

SELF-CARE*:ME
 SELF-HELP-GROUPS:ME
 (SELF next CARE)
 (SELF next MANAG*)
 (SELF next MONITOR*)
 (SELF next HELP)
 (SELF next TREAT*)
 (SELF next ADMINISTER*)
 (SELF next MEDICAT*)
 (SELF next DIAGNOS*)
 (GROUP* next SUPPORT*)
 (SUPPORT* next GROUP*)
 (PEER next SUPPORT*)
 (EXPERT next PATIENT*)
 (((PHARMACIST near SUPPORT*) or (PHARMACY near SUPPORT*)) OR
 (PHARMACIES NEAR SUPPORT*))
 (((PHARMACIST near ASSIST*) or (PHARMACY near ASSIST*)) OR
 (PHARMACIES NEAR ASSIST*))
 (((PHARMACIST near ADVISE*) or (PHARMACY near ADVISE*)) OR (PHARMACIES NEAR ADVISE*))
 (((PHARMACIST near ADVICE) or (PHARMACY near ADVICE)) OR
 (PHARMACIES NEAR ADVICE))
 (PHARMACEUTICAL next CARE)

((((((((((((((#1 or #2) or #4) or #5) or #6) or #7) or #8) or #9) or #10) or #11) or #13)
 or #14) or #15) or #16) or #17) or #18 or #19)
 ECONOMICS:ME
 COSTS-AND-COST-ANALYSIS*:ME
 ECONOMIC-VALUE-OF-LIFE*:ME
 ECONOMICS-DENTAL:ME
 ECONOMICS-HOSPITAL*:ME
 ECONOMICS-NURSING:ME
 ECONOMICS-PHARMACEUTICAL:ME
 (#21 or #22 or #23 or #24 or #25 or #26 or #27)
 (((((((((ECONOM* or COST) or COSTS) or COSTLY) or COSTING) or PRICE) or
 PRICES) or PRICING) or PHARMACOECONOMIC*)
 (EXPENDITURE* or BUDGET*)
 (VALUE next MONEY)
 (#29 or #30 or #31)
 (#28 or #32)
 (#20 and #33)

**National Research Register Issue 1 2007 (373 records/ 345 after de-
duplication)**

Accessed via Internet
 Searched 17.04.07
 Not possible to limit by date

- #1. SELF CARE explode tree 1 (MeSH) 397
- #2. SELF-HELP GROUPS single term (MeSH) 61
- #3. (self next care) 486
- #4. (self next manag*) 370
- #5. (self next monitor*) 98
- #6. (self next help) 348
- #7. (self next treat*) 27
- #8. (self next administer*) 362
- #9. (self next medicat*) 34
- #10. (self next diagnos*) 9
- #11. (group* next support*) 43
- #12. (support* next group*) 321
- #13. (peer next support*) 42
- #14. (expert next patient*) 94
- #15. (((pharmacist near support*) or (pharmacy near support*)) or (pharmacies near support*)) 43
- #16. (((pharmacist near assist*) or (pharmacy near assist*)) or (pharmacies near assist*)) 12
- #17. (((pharmacist near advise*) or (pharmacy near advise*)) or (pharmacies near advise*)) 9
- #18. (((pharmacist near advice) or (pharmacy near advice)) or (pharmacies near advice)) 15
- #19. (pharmaceutical next care) 60
- #20. (#1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19) 2028

- #21. ECONOMICS explode all trees (MeSH) 2546
- #22. COSTS AND COST ANALYSIS explode all trees (MeSH) 2156
- #23. ECONOMIC VALUE OF LIFE explode all trees (MeSH) 0
- #24. ECONOMICS DENTAL explode all trees (MeSH) 2
- #25. ECONOMICS HOSPITAL explode all trees (MeSH) 30
- #26. ECONOMICS NURSING explode all trees (MeSH) 3
- #27. ECONOMICS PHARMACEUTICAL explode all trees (MeSH) 3
- #28. (#21 or #22 or #23 or #24 or #25 or #26 or #27) 2546
- #29. (((((((econom* or cost) or costs) or costly) or costing) or price) or prices) or pricing) or pharmacoeconomic*) 10459
- #30. (expenditure* or budget*) 815
- #31. (value next money) 77
- #32. (#29 or #30 or #31) 11104
- #33. (#28 or #32) 11375
- #34. (#20 and #33) 373

HEED (Health Economic Evaluations Database) 1995-2003/Feb

Accessed via CDROM.

Searched 22.4.03

HEED (Health Economic Evaluations Database) 2003-2005/Aug (190 records)

Accessed via CDROM.

Searched 31.08.05

Ax= 'self care'
 Ax= 'self help'
 Ax= 'self manage'
 Ax= 'self managing'
 Ax= 'self management'
 Ax= 'self monitor'
 Ax= 'self monitoring'
 Ax= 'self help'
 Ax= 'self treat'
 Ax= 'self treatment'
 Ax= 'self administer'
 Ax= 'self administration'
 Ax= 'self medicate'
 Ax= 'self medication'
 Ax= 'self diagnose'
 Ax= 'self diagnosis'
 Ax= 'group support'
 Ax= 'support group'
 Ax= 'peer support'
 Ax= 'expert patient'
 Ax= 'pharmacist support'
 Ax= 'pharmacist advice'
 Ax= ' pharmacy'
 Ax= pharmacies
 Ax= 'pharmaceutical care'

HEED (Health Economic Evaluations Database) 2005-2007/April (131 records/ 113 after de-duplication)

Accessed via CDROM (April 2007).

Searched 17/04/07

Date restriction limited to year only.

Ax= 'self care'

Ax= 'self help'

Ax= 'self manage'

Ax= 'self managing'

Ax= 'self management'

Ax= 'self monitor'

Ax= 'self monitoring'

Ax= 'self help'

Ax= 'self treat'

Ax= 'self treatment'

Ax= 'self administer'

Ax= 'self administration'

Ax= 'self medicate'

Ax= 'self medication'

Ax= 'self diagnose'

Ax= 'self diagnosis'

Ax= 'group support'

Ax= 'support group'

Ax= 'peer support'

Ax= 'expert patient'

Ax= 'pharmacist support'

Ax= 'pharmacist advice'

Ax= ' pharmacy

Ax= pharmacies

Ax= 'pharmaceutical care'

Cs=1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16
or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25

JD=2005

JD=2006

JD=2007

Cs=27 or 28 or 29

Cs=26 and 30

National Guidelines Clearinghouse

Accessed via <http://www.guideline.gov/index.asp>

Searched 17.4.03

National Guidelines Clearinghouse (119 records)

Accessed via <http://www.guideline.gov/index.asp>

Searched 31.08.05

Saved as NGC.doc

"self care" "self manag*" "self help" "self treat*" "self administer*" "self medicat*" "self diagnos*" "group support*" "peer support*" "expert patient*" "pharmaceutical care"

National Guidelines Clearinghouse (105 records)

Accessed via <http://www.guideline.gov/search/detailedsearch.aspx>

Searched 18.04.07

Saved as NGC180407.doc

Keyword: "self care" or "self manag*" or "self help" or "self treat*" or "self administer*" or "self medicat*" or "self diagnos*" or "group support*" or "peer support*" or "expert patient*" or "pharmaceutical care"

Publication Date(s): 2007, 2006, 2005

National Coordinating Centre for Health Technology Assessment

Accessed via <http://www.hta.nhsweb.nhs.uk/>

Searched 17.4.03

2003-2005 update - searched as part of the HTA Database.

self care, self manage, self help, self treat, self administer, self medicat, self diagnos, group support, peer support, expert patient, pharmacist, pharmacy, pharmacies, pharmaceutical care

2005-2007 update - searched as part of the HTA Database.

self care, self manage, self help, self treat, self administer, self medicat, self diagnos, group support, peer support, expert patient, pharmacist, pharmacy, pharmacies, pharmaceutical care

Health Services Technology Assessment Text (HSTAT)

Accessed via <http://hstat.nlm.nih.gov>

Searched 22.4.03

Health Services Technology Assessment Text (HSTAT) (30 records)

Accessed via <http://hstat.nlm.nih.gov>

Searched 31.08.05

Limited to titles only. Limit by date not available

Saved as HSTAT.doc

"self care" "self help" "self mang*" "self monitor*" "self treat*" "self administer*" "self medicat*" "group support*" "expert patient*" "peer support*" "pharmacy support*" "pharmacist support*" "pharmacy advice" "pharmacist advice" "pharmaceutical care"

Health Services Technology Assessment Text (HSTAT) (34 records)

Accessed via <http://hstat.nlm.nih.gov>

Searched 18.04.07

Limited to titles only. Limit by date not available

Saved as HSTAT180407.doc

("self care"or"self help"or"self management"or"self monitoring"or"self treatment"or"self administered"or"self medicated"or"group support"or"expert patient"or"peer support"or"pharmacy support" or"pharmacist support"or"pharmacy advice"or"pharmacist advice"or"pharmaceutical care") [TITLE] AND hstat[book]

SIGN Guidelines (Scottish Intercollegiate Guidelines Network)

Accessed via <http://www.sign.ac.uk/>
Searched 23.4.03

SIGN Guidelines (Scottish Intercollegiate Guidelines Network) (0 records)

Accessed via <http://www.sign.ac.uk/>
Searched 31.08.05
Visual inspection of publications

SIGN Guidelines (Scottish Intercollegiate Guidelines Network) (0 records)

Accessed via <http://www.sign.ac.uk/>
Searched 18.04.07
Visual inspection of publications

National Institute of Clinical Excellence (published appraisals)

Accessed via <http://www.nice.org.uk/>
Searched 22.4.03

2003-2005 update - searched as part of the HTA Database.

Self care, self help, self manag, self treat, self administer, self medicate, self diagnose, group support, peer support, expert patient, pharmacy, pharmacist, pharmacies, pharmaceutical care

2005-2007 update - searched as part of the HTA Database.

Self care, self help, self manag, self treat, self administer, self medicate, self diagnose, group support, peer support, expert patient, pharmacy, pharmacist, pharmacies, pharmaceutical care

Campbell Collaboration C2 Registered Systematic Reviews titles listing

Accessed via <http://www.campbellcollaboration.org/Fralibrary.html>
Searched 23.4.03

Campbell Collaboration C2 Registered Systematic Reviews titles listing (0 records)

Accessed via <http://www.campbellcollaboration.org/Fralibrary.html>
Searched 31.08.05

No search engine. Visual inspection

Campbell Collaboration C2 Registered Systematic Reviews titles listing (0 records)

Accessed via http://www.campbellcollaboration.org/frontend.asp#Viewing_reviews

Searched 18.04.07

No search engine. Visual inspection

MEDLINE (OVIDWEB)

Accessed via <http://gateway1.uk.ovid.com/ovidweb>

Searched 13.5.03 (1966-1994)

MEDLINE (OVIDWEB)

Accessed via <http://gateway1.uk.ovid.com/ovidweb>

Searched 05.09.05 (2003-Aug 2005)

621 Records Found

- 1 self care/ or blood glucose self-monitoring/ or self administration/
- 2 self help groups/
- 3 self manag\$.ti,ab.
- 4 self care.ti,ab.
- 5 self monitor\$.ti,ab.
- 6 self help.ti,ab.
- 7 self treat\$.ti,ab.
- 8 ((self administer\$ adj2 questionnaire\$) or (self administer\$ adj2 survey\$) or (self administer\$ adj2 interview\$)).ti,ab.
- 9 self administer\$.ti,ab.
- 10 9 not 8
- 11 self medicat\$.ti,ab.
- 12 self diagnos\$.ti,ab.
- 13 (group\$ adj support\$).ti,ab.
- 14 (peer adj support\$).ti,ab.
- 15 expert patient\$.ti,ab.
- 16 ((pharmacist\$ or pharmacy or pharmacies) adj2 support\$).ti,ab.
- 17 ((pharmacist\$ or pharmacy or pharmacies) adj assist\$).ti,ab.
- 18 ((pharmacist or pharmacy or pharmacies) adj2 (advice or advise)).ti,ab.
- 19 pharmaceutical care.ti,ab.
- 20 or/1-19
- 21 economics/
- 22 exp "Costs and Cost Analysis"/
- 23 Value of Life/
- 24 economics, dental/
- 25 exp economics, hospital/
- 26 economics, medical/
- 27 economics, nursing/
- 28 economics, pharmaceutical/
- 29 or/21-28

30 (econom\$ or cost or costs or costly or costing or price or prices or pricing or
 pharmaco-economic\$.tw.
 31 (expenditure\$ not energy).tw.
 32 (value adj1 money).tw.
 33 budget\$.tw.
 34 or/30-33
 35 29 or 34
 36 letter.pt.
 37 editorial.pt.
 38 historical article.pt.
 39 or/36-38
 40 35 not 39
 41 animal/
 42 human/
 43 41 not (41 and 42)
 44 40 not 43
 45 (metabolic adj cost).ti,ab,sh.
 46 ((energy or oxygen) adj cost).ti,ab,sh.
 47 44 not (45 or 46)
 48 47 and 20

**MEDLINE & MEDLINE IN-PROCESS (OVIDWEB) (638 records/ 516 after de-
duplication)**

Accessed via <http://gateway1.uk.ovid.com/ovidweb>

Searched 17.04.07 (2005-Apr 16 2007)

638 Records Found

1 self care/ or blood glucose self-monitoring/ or self administration/ (21792)
 2 self help groups/ (5852)
 3 self manag\$.ti,ab. (3162)
 4 self care.ti,ab. (5611)
 5 self monitor\$.ti,ab. (2409)
 6 self help.ti,ab. (3002)
 7 self treat\$.ti,ab. (721)
 8 self administer\$.ti,ab. (12234)
 9 self medicat\$.ti,ab. (1607)
 10 self diagnos\$.ti,ab. (269)
 11 (group\$ adj support\$.ti,ab. (703)
 12 (peer adj support\$.ti,ab. (586)
 13 expert patient\$.ti,ab. (60)
 14 ((pharmacist\$ or pharmacy or pharmacies) adj2 support\$.ti,ab. (176)
 15 ((pharmacist\$ or pharmacy or pharmacies) adj assist\$.ti,ab. (103)
 16 ((pharmacist or pharmacy or pharmacies) adj2 (advice or advise)).ti,ab. (28)
 17 pharmaceutical care.ti,ab. (809)
 18 or/1-17 (49441)
 19 ((self administer\$ adj2 questionnaire\$) or (self administer\$ adj2 survey\$) or
 (self administer\$ adj2 interview\$)).ti,ab. (7294)
 20 18 not 19 (42147)
 21 economics/ (24856)

22 exp "Costs and Cost Analysis"/ (128397)
 23 Value of Life/ (4836)
 24 economics, dental/ (1678)
 25 exp economics, hospital/ (14634)
 26 economics, medical/ (6713)
 27 economics, nursing/ (3736)
 28 economics, pharmaceutical/ (1755)
 29 (econom\$ or cost or costs or costly or costing or price or prices or pricing or
 pharmaco-economic\$.tw. (252457)
 30 (expenditure\$ not energy).tw. (10562)
 31 (value adj1 money).tw. (11)
 32 budget\$.tw. (11190)
 33 or/21-32 (356662)
 34 letter.pt. (596264)
 35 editorial.pt. (206441)
 36 historical article.pt. (237816)
 37 or/34-36 (1031017)
 38 33 not 37 (337012)
 39 animal/ (4050286)
 40 human/ (9679026)
 41 39 not (39 and 40) (3073984)
 42 38 not 41 (320392)
 43 (metabolic adj cost).ti,ab,sh. (439)
 44 ((energy or oxygen) adj cost).ti,ab,sh. (1937)
 45 42 not (43 or 44) (318567)
 46 20 and 45 (3248)
 47 limit 46 to yr="2005 - 2007" (638)

EMBASE (OVIDWEB)

Accessed via <http://gateway1.uk.ovid.com/ovidweb>

Searched 05.09.05 (2003-Aug 2005)

776 Records Found

1. self care/ or blood glucose monitoring/ or drug self administration/
2. self medication/
3. self manag\$.ti,ab.
4. self care.ti,ab.
5. self monitor\$.ti,ab.
6. self help.ti,ab.
7. self treat\$.ti,ab.
8. ((self administer\$ adj2 questionnaire\$) or (self administer\$ adj2 survey\$) or
 (self administer\$ adj2 interview\$)).ti,ab.
9. self administer\$.ti,ab.
10. 9 not 8
11. self medicat\$.ti,ab.
12. self diagnos\$.ti,ab.
13. (group\$ adj support\$).ti,ab.
14. (peer adj support\$).ti,ab.
15. expert patient\$.ti,ab.

16. ((pharmacist\$ or pharmacy or pharmacies) adj2 support\$).ti,ab.
17. ((pharmacist\$ or pharmacy or pharmacies) adj assist\$).ti,ab.
18. ((pharmacist or pharmacy or pharmacies) adj2 (advice or advise)).ti,ab.
19. pharmaceutical care.ti,ab.
20. or/1-19
21. Health Economics/
22. exp Economic Evaluation/
23. exp Health Care Cost/
24. exp PHARMACOECONOMICS/
25. or/21-24
26. (econom\$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic\$).ti,ab.
27. (expenditure\$ not energy).ti,ab.
28. (value adj2 money).ti,ab.
29. budget\$.ti,ab.
30. or/26-29
31. 25 or 30
32. (metabolic adj cost).ti,ab.
33. ((energy or oxygen) adj cost).ti,ab.
34. ((energy or oxygen) adj expenditure).ti,ab.
35. or/32-34
36. 31 not 35
37. editorial.pt.
38. note.pt.
39. letter.pt.
40. or/37-39
41. 36 not 40
42. (rat or rats or mouse or mice or hamster or hamsters or animal or animals or dogs or dog or cats or bovine or sheep).ti,ab,sh.
43. exp animal/
44. Nonhuman/
45. or/42-44
46. exp human/
47. exp human experiment/
48. 46 or 47
49. 45 and (45 not 48)
50. 41 not 49
51. 20 and 50

EMBASE (OVIDWEB) (692 records/ 369 after de-deuplication)

Accessed via <http://gateway1.uk.ovid.com/ovidweb>

Searched 17.04.07 (2005-2007 Week 15)

692 Records Found

- 1 self care/ or blood glucose monitoring/ or drug self administration/ (11656)
- 2 self medication/ (4355)
- 3 self manag\$.ti,ab. (2515)
- 4 self care.ti,ab. (3254)
- 5 self monitor\$.ti,ab. (2164)

6 self help.ti,ab. (2487)
7 self treat\$.ti,ab. (649)
8 self administer\$.ti,ab. (10310)
9 self medicat\$.ti,ab. (1583)
10 self diagnos\$.ti,ab. (202)
11 (group\$ adj support\$).ti,ab. (549)
12 (peer adj support\$).ti,ab. (400)
13 expert patient\$.ti,ab. (49)
14 ((pharmacist\$ or pharmacy or pharmacies) adj2 support\$).ti,ab. (201)
15 ((pharmacist\$ or pharmacy or pharmacies) adj assist\$).ti,ab. (108)
16 ((pharmacist or pharmacy or pharmacies) adj2 (advice or advise)).ti,ab. (44)
17 pharmaceutical care.ti,ab. (1370)
18 or/1-17 (34898)
19 ((self administer\$ adj2 questionnaire\$) or (self administer\$ adj2 survey\$) or
(self administer\$ adj2 interview\$)).ti,ab. (5918)
20 18 not 19 (28980)
21 health economics/ (9122)
22 exp economic evaluation/ (86465)
23 exp health care cost/ (87889)
24 exp pharmacoeconomics/ (45594)
25 (econom\$ or cost or costs or costly or costing or price or prices or pricing or
pharmacoeconomic\$).tw. (197375)
26 (expenditure\$ not energy).tw. (8408)
27 (value adj1 money).tw. (7)
28 budget\$.tw. (7794)
29 or/21-28 (290198)
30 (metabolic adj cost).ti,ab. (340)
31 ((energy or oxygen) adj cost).ti,ab. (1566)
32 ((energy or oxygen) adj expenditure).ti,ab. (8747)
33 or/30-32 (10182)
34 29 not 33 (287887)
35 note.pt. (209771)
36 letter.pt. (344831)
37 editorial.pt. (176176)
38 or/35-37 (730778)
39 34 not 38 (249762)
40 (rat or rats or mouse or mice or hamster or hamsters or animal or animals or
dogs or dog or cats or bovine or sheep).ti,ab,sh. (1851048)
41 exp animal/ (112185)
42 nonhuman/ (2867095)
43 or/40-42 (3184600)
44 exp human/ (5699081)
45 exp human experiment/ (231763)
46 or/44-45 (5699944)
47 43 and (43 not 46) (2645875)
48 39 not 47 (229916)
49 20 and 48 (3004)
50 limit 49 to yr="2005 - 2007" (692)

Appendix B

Checklist used for quality assessment of economic evaluations

Study specification

1. Was the study question clear?

- ☐ What was the objective of the study? How clear was it? Self-management interventions (SC) can be complex.
- ☐ What were the key costs/resources and outcomes assessed?

2. Was a comprehensive description of the competing alternatives given?

- ☐ Can you tell who? did what? to whom? where? and how often? Were the care pathways clearly identified
- ☐ Was a comprehensive description of the interventions provided (i.e. the delivery needs to be described in detail in order to standardise the delivery and to generalise the results)
- ☐ Were any alternatives omitted? Is it appropriate to have a do-nothing alternative; could the study design prevent people self medicating?
- ☐ For SC interventions given in groups, was the effect of group therapy allowed for (e.g. by inclusion of the comparator arm of group therapy without the intervention). What was the unit of analysis?
- ☐ Were baseline utility values and/or resource use given for alternative treatment arms?

3. What was the perspective of the study?

- ☐ Often economic studies use a health service perspective on the grounds that this approximates a societal perspective. In the case of SC evaluations, costs to patient, plus productivity changes should be explored.

4. What was the study design?

- ☐ RCT of comparator therapies, placebo controlled RCT, controlled before and after, cross sectional control, before after control, case reports, expert opinion.

5. What was the economic study type?

- ☐ Cost utility, cost-effectiveness, cost benefit, cost minimisation, cost consequences
-

Clinical evidence

6. Given the type of study was the design adequate?

Sample.

- ☐ How was the sample size determined?
- ☐ Was sample size adequate to detect differences? No statistically significant differences may lead to inappropriate cost minimisation analysis.
- ☐ How was the sample selected?
- ☐ Is there evidence to suggest that the sample is appropriate for the study question? Did any subjects refuse to participate?
- ☐ Were there any in/exclusions criteria and if so were they appropriate?
- ☐ Did studies which were RCTs have proper randomisation process?

Contamination.

- ☐ Given that contamination of control groups is more likely in SC interventions, was contamination considered? Were attempts made to mitigate it?
- ☐ Was adherence/compliance with the intervention measured?

Sociodemographic characteristics of study populations.

- ☐ Were these reported?
- ☐ Were the study and control groups shown to be comparable at baseline in terms of socio-demographic characteristics?
- ☐ Factors such as chronicity, previous treatment, social adjustment, interpersonal difficulties and social circumstances may also impact on the outcomes

Compliance. Was compliance reported?

Is the setting described?

- ☐ Is the area/country identified?
- ☐ Where did the intervention take place (home, primary care, hospital...)
- ☐ Who delivered the intervention?
- ☐ How many centres were there?
- ☐ Was effectiveness established in a UK trial?

Dates.

- ☐ Was the date of the intervention given?
- ☐ Were the dates for the effectiveness measures, resource use and price given?

Outcome assessment

- ☐ Is the method described?
- ☐ Was assessment blinded?

Economic analysis

7. Were all the important and relevant costs and consequences for each alternative identified?

7a Costs

- ☐ Was the costing undertaken on the same sample of people as that used in the effectiveness study?
- ☐ Was the costing undertaken pro/retrospectively?
- ☐ Self-management interventions may have more influence on self medication; with these interventions, these costs should be included.
- ☐ Formal and informal care costs may also vary more markedly in SC interventions as the person becomes better able to cope for him/herself and should be measured

7b Consequences

- ☐ SC interventions have been associated with improvements in outcomes such as empowerment, confidence etc.
- ☐ Were any adverse effects reported?
- ☐ Was quality of life measured?
- ☐ Appropriateness of outcome measures?
- ☐ Were patient preferences explored?

8. Were costs and consequences measured accurately (and credibly) in appropriate physical units? (e.g. hours of nursing time, number of physician visits, lost work-days, gained life-years)

8a. Costs

- ☐ Where were unit cost data derived from? Were they derived from the study? Were they UK based? If not, are these estimates good estimates of opportunity cost in UK?
- ☐ Was study powered on costs?
- ☐ What resource quantities and costs were reported and were they reported separately and how were they estimated?
- ☐ What direct costs were included?
- ☐ What productivity changes were included? Were working days really lost?
- ☐ If required, were appropriate adjustments for inflation/currency conversion made?
- ☐ Prospective data for the study? From elsewhere?

8b. Consequences

- ☐ Use of outcome measures such as empowerment, motivation, perseverance. Have the measures been validated and are they reliable? Whose values were used and how many?
- ☐ Was the follow up period adequate? Was there loss to follow up? Was the outcome analysis built on intention to treat/treatment completers? Were the outcomes assessable within the timeframe? SC interventions may have little immediate impact on “hard” outcomes (mortality, life years gained etc);. If short follow up, is the link between intermediate outcome (confidence etc) and final outcome (e.g. QALY), well established?
- ☐ Direct or indirect measures of health effect? Health utility analysis?

Analysis

9. Was the statistical analysis appropriate given the design?

- ☐ EG allowance for clustering, loss to follow up
- 10a Was sub-group analysis performed?**
- 10b If so were the groups pre-specified?**
 - ☐ SC interventions often study a very heterogeneous population e.g. with number of conditions, age group.
- 11. Were costs and consequences appropriately discounted?**
 - ☐ Were discount factors applied to costs and outcomes if appropriate?
 - ☐ Was a lower discount rate applied to outcomes if these were measured in volume terms rather than in value terms?
- 12. Was an incremental analysis of costs and consequences of alternatives performed?**
 - ☐ Were the additional (incremental) costs generated by one alternative over another compared to the additional effects, benefits or utilities generated?
- 13. Was allowance made for uncertainty?**
 - Stochastic analysis of patient-level data**
 - ☐ Were details of statistical tests and confidence intervals given for stochastic data?
 - ☐ Was uncertainty around cost-effectiveness expressed (e.g. confidence interval around incremental cost-effectiveness ratio (ICER), cost-effectiveness acceptability curves).
 - ☐ Was sensitivity analysis used to assess uncertainty in non-stochastic variables (e.g. unit costs, discount rates) and analytic decisions (e.g. methods to handle missing data)?
 - Stochastic analysis of decision models**
 - ☐ Are all appropriate input parameters included with uncertainty?
 - ☐ Is second-order uncertainty (uncertainty in means) included rather than first order (uncertainty between patients)?
 - ☐ Are the probability distributions adequately detailed and appropriate? Was sensitivity analysis used to assess uncertainty in non-stochastic variables (e.g. unit costs, discount rates) and analytic decisions (e.g. methods to handle missing data)?
 - Deterministic analysis**
 - ☐ Was the method of sensitivity analysis used specified and justified (e.g. univariate, threshold analysis etc)?
 - ☐ Was the choice of variables used in the sensitivity analysis justified and the ranges over which the variables are varied stated?
- 14. Were missing data handled adequately?**
 - ☐ As with most economic evaluations, missing data are likely to be a problem with SC interventions.

15a. Was an economic model developed?

15b. If so was it appropriate, transparent and the methodology explicit?

16 Study results and implications

16a. Were limitations of the study acknowledged and biases (and their potential magnitude and direction) discussed?

- ☐ sources of funding acknowledged; declarations of competing interests

16b. Were issues of generalisability discussed?

- ☐ different country, different institutional setting, standard intervention, population groups.

16c. Did the presentation and discussion of study results include all important issues?

- ☐ Were any recommendations made by the authors regarding policy/practice?
 - ☐ Were specific recommendations made by the authors regarding the need for further research?
-

Appendix C

C.1 R code for IPD meta-analysis in Chapter 5

```
rm(list=ls(all=TRUE))
library("R2WinBUGS")
library("BRugs")
setwd('C:/work/phD_stuff/')
tData <- read.table(file='QALY5.TXT',header=TRUE)

N.SIMS <- 15000
N.BURNIN <-5000
WIN.DEBUG <- FALSE
N.CHAINS <- 2

WIN.SOURCES <- c('GR7NSHM1Norm.txt','GR7NSHM1Log.txt','GR7NSHM1LogMult.txt')
#WIN.SOURCES <- c('GR7NSHM1Log.txt')
nWinSource <- length(WIN.SOURCES)

genSamps <- function(data,
                      n.iter=N.SIMS+N.BURNIN,
                      inits,n.burnin=N.BURNIN,
                      bugs.file=winSource,
                      parameters.to.save,
                      winSource=winSource,
                      winDebug=WIN.DEBUG){
  sims <- bugs(data=data,
               parameters.to.save=parameters.to.save,
               model.file=bugs.file,
               n.chains=N.CHAINS,
               debug=winDebug,
               inits=inits,
               n.iter = n.iter,
               n.burnin = n.burnin,
               n.thin=1,
               program = "openbugs"
               )
  return(sims)
}

qaly <- tData$QALY
cost <- tData$cost
tx <- tData$tx
study <- tData$trial_id
nTx<-length(unique(tx))
nObs <-nrow(tData)
nStudies <- max(study)+1

age <- tData$age
age <- age-mean(age,na.rm=TRUE)
gender<-1+tData$gender
EQ-5D<-tData$base_EQ-5D
EQ-5D<-EQ-5D-mean(EQ-5D)
se1 <- tData$selfeff1
```



```

se1 <- se1-mean(se1,na.rm=TRUE)
se2<-tData$se2
se2 <- se2-mean(se2,na.rm=TRUE)
qalyMean <- matrix(NA,nStudies-1,nTx)
costMean <- matrix(NA,nStudies-1,nTx)
for (jj in 1:nTx){
  for (ss in 1:(nStudies-1)){
    qalyMean[ss,jj] <- mean(qaly[tx==jj-1&study==ss],na.rm=TRUE)
    costMean[ss,jj] <- mean(cost[tx==jj-1&study==ss],na.rm=TRUE)
  }
}

qalyMean[is.na(qalyMean)] <- 0
costMean[is.na(qalyMean)] <- 0

studyDummy <- study-1
#maybe se1, note that lm rather than glm generates OLS estimates
summary(glm(qaly~tx+EQ-5D+studyDummy))
summary(glm(cost~tx+EQ-5D+studyDummy,gaussian(link = "log")))
summary(lm(cost~tx+EQ-5D+studyDummy))

#analyse effic data
inits <- function(){list(qSe=rep(1,nStudies),logCostSe=rep(1,nStudies),
                        ctBeta=1,qtBeta=1,df=1,ceBeta=rnorm(nStudies),qeBeta=rnorm(nStudies),
                        cAlpha=rep(1000,nStudies),qAlpha=rep(1,nStudies))}

bugsData <- list(cost=cost,qaly=qaly,qSampMean=qalyMean,nObs=nObs,tx=tx,
                nTx=nTx,study=study,EQ-5D=EQ-5D,nStudies=nStudies)

parameters.to.save <-
c("ctBeta","qtBeta","cSe","qSe","cAlpha","deltaCost","deltaQaly","postPredC","postPredQ","r
esidCost","residQaly")
#parameters.to.save <- c("deltaCost","deltaQaly","df")

#model 1 basic model with no covs or corr, normally distributed costs if norm suffix
for (ww in 1:nWinSource){

res1 <-
genSamps(data=bugsData,inits=inits,parameters.to.save=parameters.to.save,winSource=
WIN.SOURCES[ww])
nonTxcost1 <- res1$sims.matrix["deltaCost"]
qaly1 <- res1$sims.matrix["deltaQaly"]
ctBeta1 <- res1$sims.matrix["ctBeta"]
qtBeta1 <- res1$sims.matrix["qtBeta"]
txCost <- 198
cost1 <- nonTxcost1+txCost
ceRange <- seq(0,100000,100)
ceac <- numeric()
icer<-numeric()

for (cc in seq(along=ceRange)){
  lambda <- ceRange[cc]
  nb1<-qaly1*lambda-cost1
  ceac[cc] <- mean(nb1>0)
  nb1[nb1<0]<-0

```

```

    #nb2[nb<0]<-0
    #modelVoi[cc,ii] <- mean(nb)-mean(seNb)
  }

results <- rbind(
  c(mean(nonTxcost1),quantile(nonTxcost1,c(0.025,0.975))),
  c(mean(qaly1),quantile(qaly1,c(0.025,0.975))),
  c(mean(ctBeta1),quantile(ctBeta1,c(0.025,0.975))),
  c(mean(qtBeta1),quantile(qtBeta1,c(0.025,0.975))))
results <- rbind(results,results[1,]/results[2,])
colnames(results) <- c("Mean","2.5%","97.5%")
rownames(results) <- c("Cost","QALY","BCT","BQT","ICER")
print(WIN.SOURCES[ww])
print(results)
costResid <- res1$summary[(14+nObs*2):(14+nObs*3-1),1]
qalyResid <- res1$summary[(14+nObs*3):(14+nObs*4-1),1]
costPValues <- (apply(res1$sims.matrix[,14:(14+nObs-1)]>0,2,mean))
qalyPValues <- (apply(res1$sims.matrix[, (14+nObs):(14+2*nObs-1)]>0,2,mean))
DIC <- res1$DIC
print(DIC)

assign(paste("dic",ww,sep=""),DIC)
assign(paste("results",ww,sep=""),results)
assign(paste("qalyPValues",ww,sep=""),qalyPValues)
assign(paste("costPValues",ww,sep=""),costPValues)
assign(paste("qalyResid",ww,sep=""),qalyResid)
assign(paste("costResid",ww,sep=""),costResid)
assign(paste("ceac",ww,sep=""),ceac)

remove(res1)
remove(qaly1)
remove(ctBeta1)
remove(qtBeta1)
}

rm(list=ls(all=TRUE))
library("R2WinBUGS")
library("BRugs")
setwd('C:/work/phD_stuff/neil3')
tData <- read.table(file='QALY5.TXT',header=TRUE)

#tData <- tData[tData$trial_id==1,]
#tData <- tData[tData$trial_id==2,]

N.SIMS <- 15000
N.BURNIN <- 5000
WIN.DEBUG <- FALSE
N.CHAINS <- 2

WIN.SOURCES <-
c('GR7NSHM1Norm.txt','GR7NSHM1Log.txt','GR7NSHM1LogMult.txt','GR7NSHM1TDist.txt'
)
#WIN.SOURCES <- c('GR7NSHM1Log.txt')
nWinSource <- length(WIN.SOURCES)

```



```

genSamps <- function(data,
                     n.iter=N.SIMS+N.BURNIN,
                     inits,n.burnin=N.BURNIN,
                     bugs.file=winSource,
                     parameters.to.save,
                     winSource=winSource,
                     winDebug=WIN.DEBUG){
  sims <- bugs(data=data,
               parameters.to.save=parameters.to.save,
               model.file=bugs.file,
               n.chains=N.CHAINS,
               debug=winDebug,
               inits=inits,
               n.iter = n.iter,
               n.burnin = n.burnin,
               n.thin=1,
               program = "openbugs"
             )
  return(sims)
}

qaly <- tData$QALY
cost <- tData$cost
tx <- tData$tx
study <- tData$trial_id
nTx<-length(unique(tx))
nObs <-nrow(tData)
nStudies <- max(study)+1
age <- tData$age
age <- age-mean(age,na.rm=TRUE)
gender<-1+tData$gender
EQ-5D<-tData$base_EQ-5D
EQ-5D<-EQ-5D-mean(EQ-5D)
se1 <- tData$selfeff1
se1 <- se1-mean(se1,na.rm=TRUE)
se2<-tData$se2
se2 <- se2-mean(se2,na.rm=TRUE)
qalyMean <- matrix(NA,nStudies-1,nTx)
costMean <- matrix(NA,nStudies-1,nTx)

for (jj in 1:nTx){
  for (ss in 1:(nStudies-1)){
    qalyMean[ss,jj] <- mean(qaly[tx==jj-1&study==ss],na.rm=TRUE)
    costMean[ss,jj] <- mean(cost[tx==jj-1&study==ss],na.rm=TRUE)
  }
}
qalyMean[is.na(qalyMean)] <- 0
costMean[is.na(qalyMean)] <- 0
studyDummy <- study-1
#maybe se1, note that lm rather than glm generates OLS estimates
summary(glm(qaly~tx+EQ-5D+studyDummy))
summary(glm(cost~tx+EQ-5D+studyDummy,gaussian(link = "log")))
summary(lm(cost~tx+EQ-5D+studyDummy))

#analyse effic data

```

```

inits <- function(){list(qSe=rep(1,nStudies),logCostSe=rep(1,nStudies),
                        ctBeta=1,qtBeta=1,df=1,ceBeta=rnorm(nStudies),qeBeta=rnorm(nStudies),
                        cAlpha=rep(1000,nStudies),qAlpha=rep(1,nStudies))}

bugsData <- list(cost=cost,qaly=qaly,qSampMean=qalyMean,nObs=nObs,tx=tx,
                nTx=nTx,study=study,EQ-5D=EQ-5D,nStudies=nStudies)

parameters.to.save <-
c("ctBeta","qtBeta","cSe","qSe","cAlpha","deltaCost","deltaQaly","postPredC","postPredQ","r
esidCost","residQaly")
#parameters.to.save <- c("deltaCost","deltaQaly","df")

#model 1 basic model with no covs or corr, normally distributed costs if norm suffix
for (ww in 1:nWinSource[4]){

res1 <-
genSamps(data=bugsData,inits=inits,parameters.to.save=parameters.to.save,winSource=
WIN.SOURCES[ww])
nonTxcost1 <- res1$sims.matrix["deltaCost"]
qaly1 <- res1$sims.matrix["deltaQaly"]
ctBeta1 <- res1$sims.matrix["ctBeta"]
qtBeta1 <- res1$sims.matrix["qtBeta"]
txCost <- 198

cost1 <- nonTxcost1+txCost
ceRange <- seq(0,100000,100)
ceac <- numeric()
icer<-numeric()

for (cc in seq(along=ceRange)){
  lambda <- ceRange[cc]
  nb1<-qaly1*lambda-cost1
  ceac[cc] <- mean(nb1>0)
  nb1[nb1<0]<-0
  #nb2[nb<0]<-0
  #modelVoi[cc,ii] <- mean(nb)-mean(seNb)
}
results <- rbind(
  c(mean(nonTxcost1),quantile(nonTxcost1,c(0.025,0.975))),
  c(mean(qaly1),quantile(qaly1,c(0.025,0.975))),
  c(mean(ctBeta1),quantile(ctBeta1,c(0.025,0.975))),
  c(mean(qtBeta1),quantile(qtBeta1,c(0.025,0.975))))
results <- rbind(results,results[1,]/results[2,])
colnames(results) <- c("Mean","2.5%","97.5%")
rownames(results) <- c("Cost","QALY","BCT","BQT","ICER")

print(WIN.SOURCES[ww])
print(results)

costResid <- res1$summary[(14+nObs*2):(14+nObs*3-1),1]
qalyResid <- res1$summary[(14+nObs*3):(14+nObs*4-1),1]

costPValues <- (apply(res1$sims.matrix[,14:(14+nObs-1)]>0,2,mean))
qalyPValues <- (apply(res1$sims.matrix[, (14+nObs):(14+2*nObs-1)]>0,2,mean))

```



```

DIC <- res1$DIC
print(DIC)

assign(paste("dic",ww,sep=""),DIC)
assign(paste("results",ww,sep=""),results)
assign(paste("qalyPValues",ww,sep=""),qalyPValues)
assign(paste("costPValues",ww,sep=""),costPValues)
assign(paste("qalyResid",ww,sep=""),qalyResid)
assign(paste("costResid",ww,sep=""),costResid)
assign(paste("ceac",ww,sep=""),ceac)

remove(res1)
remove(qaly1)
remove(ctBeta1)
remove(qtBeta1)
}
layout(1)
plot(0,0,xlim=range(ceRange)/1000,ylim=c(0,1),xlab="Cost-effectiveness Threshold
(£1000,s)",ylab="Probability EPP
cost-effective",bty="l",type="n")
for (ww in 1:nWinSource){
  lines(ceRange/1000,get(paste("ceac",ww,sep="")),col=ww,lwd=6,lty=ww)
}
legend(locator(1),WIN.SOURCES,col=1:nWinSource,lty=1)
layout(cbind((1:nWinSource*2)-1,1:nWinSource*2))
for (ww in 1:nWinSource){
  plot(cost,get(paste("costResid",ww,sep="")),pch=19,ylab="Actual - Predicted
Cost",xlab="Cost",main=WIN.SOURCES[ww])

  plot(1:length(get(paste("costPValues",ww,sep=""))),sort(get(paste("costPValues",ww,sep="
"))),xlab=" ",ylab=" ")
}

plot(qaly,qalyResid,pch=19,ylab="Actual - Predicted Qalys",xlab="Qalys",main=winSource)

plot(1:length(qalyPValues),sort(qalyPValues),main=winSource)
layout( cbind(c(1,3),c(2,4)) )

hist(cost[study==1&tx==0],xlim=c(0,25000))
hist(cost[study==1&tx==1],xlim=c(0,25000))
hist(cost[study==2&tx==0],xlim=c(0,10000))
hist(cost[study==2&tx==1],xlim=c(0,10000))

hist(qaly[study==1&tx==0],xlim=c(0,1))
hist(qaly[study==1&tx==1],xlim=c(0,1))
hist(qaly[study==2&tx==0],xlim=c(0,1))
hist(qaly[study==2&tx==1],xlim=c(0,1))

```

C.2 WinBUGS code for IPD meta analysis in Chapter 5

C.2.1 Normally distributed costs

```

Model{
  for(ii in 1:nObs){

```

#MODEL 1: TRIAL BASED ANALYSIS

```
cMean[iii] <- cAlpha[study[iii]] + ctBeta*tx[iii] + ceBeta[study[iii]]*EQ-5D[iii]
```

```
cc[iii] <- corrBeta[study[iii],tx[iii]+1] * (qaly[iii] - qSampMean[study[iii],tx[iii]+1])
```

```
logCMean[iii] <- log(cMean[iii])-0.5*cVar[study[iii]]
```

```
cost[iii] ~ dnorm(cMean[iii],cPrec[study[iii]]) #trial c mean and precision
```

```
qaly[iii] ~ dnorm(qMean[iii], qPrec[study[iii]])
```

```
#trial q mean,precision
```

```
qMean[iii] <- qAlpha[study[iii]] + qtBeta*tx[iii] + qeBeta[study[iii]]*EQ-5D[iii]
```

```
residCost[iii] <- cost[iii] - cMean[iii]
```

```
residQaly[iii] <- qaly[iii] - qMean[iii]
```

#we have a value for each individual, the predictive posterior gives a distribution of predicted estimates given all covs etc and we #can see how the actual observed value

```
newC[iii] ~ dnorm(cMean[iii],cPrec[study[iii]])
```

```
newQ[iii] ~ dnorm(qMean[iii],qPrec[study[iii]])
```

```
postPredC[iii] <- cost[iii] - newC[iii]
```

```
postPredQ[iii] <- qaly[iii] - newQ[iii]
```

```
}
```

```
df ~ dunif(1,10)
```

```
for (ss in 1:nStudies) {
```

```
  for (tt in 1:nTx) {
```

```
    corrBeta[ss,tt] ~ dunif(-5,5)
```

```
    #psi[ss,tt] <- corrBeta[ss,tt]*sqrt(qvar[ss])/sqrt(cvar[ss])
```

```
  }
```

```
  ceBeta[ss]~ dnorm (0,1.0E-12)
```

```
  qeBeta[ss]~ dnorm (0,1.0E-12)
```

```
  cAlpha[ss] ~ dnorm (0,1.0E-12)
```

```
  qAlpha[ss] ~ dnorm (0,1.0E-12)
```

```
  seAlpha[ss] ~ dnorm (0,1.0E-12)
```

```
  qSe[ss] ~ dunif (0,5)
```

```
  qVar[ss] <- pow(qSe[ss],2)
```

```
  qPrec[ss] <- 1/qVar[ss]
```

```
  logCostSe[ss] ~ dunif (-5,10)
```

```
  cSe[ss] <- exp(logCostSe[ss])
```

```
  cVar[ss] <- pow(cSe[ss],2)
```

```
  cPrec[ss]<-1/cVar[ss]
```

```
  sepSe[ss] ~ dunif (0,10)
```

```
  sePrec[ss]<-pow(sepSe[ss], -2)
```

```
}
```



```

deltaCost <- ctBeta
deltaQaly <- qtBeta

ctBeta~ dnorm (0,1.0E-12)
qtBeta~ dnorm (0,1.0E-12)

}

C.2.2 Log normal costs

Model{
  for(ii in 1:nObs){
    #MODEL 1: TRIAL BASED ANALYSIS

    cMean[ii] <- cAlpha[study[ii]] + ctBeta*tx[ii] + ceBeta[study[ii]]*EQ-5D[ii]

    cc[ii] <- corrBeta[study[ii],tx[ii]+1] * (qaly[ii] - qSampMean[study[ii],tx[ii]+1])

    logCMean[ii] <- log(cMean[ii])-0.5*cVar[study[ii]]
    cost[ii] ~ dlnorm(logCMean[ii],cPrec[study[ii]]) #trial c mean and precision

    qaly[ii] ~ dnorm(qMean[ii], qPrec[study[ii]]) #trial q mean,precision
    qMean[ii] <- qAlpha[study[ii]] + qtBeta*tx[ii] + qeBeta[study[ii]]*EQ-5D[ii]

    residCost[ii] <- cost[ii] - cMean[ii]
    residQaly[ii] <- qaly[ii] - qMean[ii]

    #we have a value for each individual, the predictive posterior gives a distribution of predicted
    estimates given all covs etc and we #can see how the actual observed value
    newC[ii] ~ dlnorm(logCMean[ii],cPrec[study[ii]])
    newQ[ii] ~ dnorm(qMean[ii],qPrec[study[ii]])
    postPredC[ii] <- cost[ii] - newC[ii]
    postPredQ[ii] <- qaly[ii] - newQ[ii]
  }
  df ~ dunif(1,10)
  for (ss in 1:nStudies) {
    for (tt in 1:nTx) {
      corrBeta[ss,tt] ~ dunif(-5,5)
      # psi[ss,tt] <- corrBeta[ss,tt]*sqrt(qvar[ss])/sqrt(cvar[ss])
    }
    ceBeta[ss]~ dnorm (0,1.0E-12)
    qeBeta[ss]~ dnorm (0,1.0E-12)

    cAlpha[ss] ~ dnorm (0,1.0E-12)
    qAlpha[ss] ~ dnorm (0,1.0E-12)
    seAlpha[ss] ~ dnorm (0,1.0E-12)
    qSe[ss] ~ dunif (0,5)
    qVar[ss] <- pow(qSe[ss],2)
    qPrec[ss] <- 1/qVar[ss]
    logCostSe[ss] ~ dunif (-5,10)
    cSe[ss] <- exp(logCostSe[ss])
    cVar[ss] <- pow(cSe[ss],2)
    cPrec[ss]<-1/cVar[ss]

    sepSe[ss] ~ dunif (0,10)
    sePrec[ss]<-pow(sepSe[ss], -2)
  }
}

```

```
deltaCost <- ctBeta  
deltaQaly <- qtBeta  
ctBeta~ dnorm (0,1.0E-12)  
qtBeta~ dnorm (0,1.0E-12)
```


Appendix D

Search strategy for populating the models in Chapter 6

MEDLINE and MEDLINE In Process

MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R)
1950 to Present

Searched via OVID 02/07/07

- 1 exp self efficacy/ (4419)
- 2 self efficacy.ti,ab. (5014)
- 3 or/1-2 (7493)
- 4 exp quality-adjusted life year/ (2952)
- 5 "quality-adjusted life year\$.ti,ab. (2114)
- 6 "quality adjusted life year\$.ti,ab. (2114)
- 7 QALY.ti,ab. (1388)
- 8 or/4-7 (4100)
- 9 utility.ti,ab. (58358)
- 10 3 and 8 (13)
- 11 3 and 9 (133)
- 12 10 or 11 (144)
- 13 limit 12 to english language (142)
- 14 limit 13 to yr="1990 - 2007" (138)

Results file: checrd on rentedfs/CHE/gerry/self efficacy medline 138.txt

EMBASE

EMBASE 1980 to 2007 Week 26

Searched via OVID 02/07/07

- 1 self efficacy.ti,ab. (3887)
- 2 exp quality-adjusted life year/ (3063)
- 3 "quality-adjusted life year\$.ti,ab. (1974)
- 4 "quality adjusted life year\$.ti,ab. (1974)
- 5 QALY.ti,ab. (1293)
- 6 or/2-5 (3890)
- 7 utility.ti,ab. (53311)
- 8 1 and 6 (2)
- 9 1 and 7 (85)
- 10 8 or 9 (87)
- 11 limit 10 to (english language and yr="1990 - 2007") (82)

Results file: checrd on rentedfs/CHE/gerry/self efficacy embase 82.txt

HMIC

HMIC Health Management Information Consortium May 2007

Searched via OVID 02/07/07

- 1 self efficacy.ti,ab. (104)
- 2 exp quality-adjusted life years/ (153)
- 3 "quality-adjusted life year\$.ti,ab. (304)
- 4 "quality adjusted life year\$.ti,ab. (304)
- 5 QALY.ti,ab. (215)

- 6 or/2-5 (417)
- 7 utility.ti,ab. (747)
- 8 1 and 6 (1)
- 9 1 and 7 (1)
- 10 8 or 9 (2)
- 11 limit 10 to yr="1990 - 2007" (2)

Results file: checrd on rentedfs/CHE/ gerry/self efficacy hmic 2.txt

Cochrane Central Register of Clinical Trials (CENTRAL)

Searched via Wiley 02/07/07

#1 (self efficacy):ti,ab,kw 3462

#2 (utility):ti,ab,kw 2967

#3 (QALY):ti,ab,kw 172

#4 "quality adjusted life year":ti,ab,kw 129

#5 (#3 OR #4) 240

#6 (#1 AND #2) 43

#7 (#1 AND #5) 0

#8 (#6 OR #7), from 1990 to 2007 39

Of 39 results in entire Cochrane Library- 37 are from CENTRAL.

Results file: checrd on rentedfs/CHE/gerry/self efficacy central 37.txt

Results

Database	Results	After dedupe	Custom 4
MEDLINE/ MEDLINE In Process	138	137	Medline and Medline In Process 02/07/07
EMBASE	82	13	EMBASE 02/07/07
HMIC	2	0	HMIC 02/07/07
CENTRAL	37	23	CENTRAL 02/07/07
Total	259	173	

Appendix E

E.1 R code for alternative models used in Chapter 6

```
library("R2WinBUGS")
library("BRugs")

setwd('z:/gar2/')
tData <- read.table(file='QALY5.TXT',header=TRUE)

#tData <- tData[tData$trial_id==1,]

aggStudy <- c(1,2,3,4,5,6,7,8)
aggSe2 <- c(0.5,0.310,1.41,0.57,0.63,0.44,1.14,0.206)
aggSeSD <- c(2.4,1.67,2.53,3.45,3.81,4.36,3.11,2.69)
aggN <- c(489,430,177,148,954,327,286,780)
aggNobs <- length(aggStudy)

N.SIMS <- 15000
N.BURNIN <- 5000
WIN.DEBUG <- FALSE
N.CHAINS <- 2

genSamps <- function(data,
                      n.iter=N.SIMS+N.BURNIN,
                      inits,n.burnin=N.BURNIN,
                      bugs.file=WIN.SOURCE,
                      parameters.to.save,
                      winSource=WIN.SOURCE,
                      winDebug=WIN.DEBUG){

  sims <- bugs(data=data,
               parameters.to.save=parameters.to.save,
               model.file=bugs.file,
               n.chains=N.CHAINS,
               debug=winDebug,
               inits=inits,
               n.iter = n.iter,
               n.burnin = n.burnin,
               n.thin=1,
               program='openbugs'
               )
  return(sims)
}

qaly <- tData$QALY
cost <- tData$cost
tx <- tData$tx
study <- tData$trial_id
nTx<-length(unique(tx))
```

```

nObs <- nrow(tData)
nStudies <- max(study)+1

age <- tData$age
age <- age-mean(age,na.rm=TRUE)

gender<-1+tData$gender

EQ-5D<-tData$base_EQ-5D
EQ-5D<-EQ-5D-mean(EQ-5D)

se1 <- tData$selfeff1
se1 <- se1-mean(se1,na.rm=TRUE)

se2<-tData$se2
se2 <- se2-mean(se2,na.rm=TRUE)

qalyMean <- matrix(NA,nStudies-1,nTx)
costMean <- matrix(NA,nStudies-1,nTx)

for (jj in 1:nTx){
  for (ss in 1:(nStudies-1)){
    qalyMean[ss,jj] <- mean(qaly[tx==jj-1&study==ss],na.rm=TRUE)
    costMean[ss,jj] <- mean(cost[tx==jj-1&study==ss],na.rm=TRUE)
  }
}

qalyMean[is.na(qalyMean)] <- 0
costMean[is.na(qalyMean)] <- 0

studyDummy <- study-1
#maybe se1
qalyReg <- summary(glm(qaly~tx+EQ-5D*studyDummy+studyDummy))
costReg <- summary(glm(cost~tx+EQ-5D*studyDummy+studyDummy))
icer <- (costReg$coeff[2,1]+250)/qalyReg$coeff[2,1]

#analyse effic data
inits <- function(){list(qSe=rep(1,nStudies),logCostSe=rep(1,nStudies), csBeta=0,qsBeta=0,
s1Alpha=rep(1,nStudies),s2Alpha=rep(1,nStudies),s1pSe=rep(1,nStudies),s2pSe=rep(1,nStudies),
ctBeta=1,qtBeta=1,ceBeta=rnorm(nStudies),qeBeta=rnorm(nStudies),s2tBeta=1,
cAlpha=rnorm(nStudies,1000,20),qAlpha=rnorm(nStudies,1,0.2))}
#analyse effic data
#inits <- function(){list(qSe=runif(nStudies,0,1),logCostSe=runif(nStudies,-1,1),
#
ctBeta=rnorm(1),qtBeta=rnorm(1),cAlpha=rnorm(nStudies),qAlpha=rnorm(nStudies))}

bugsData <- list(cost=cost,qaly=qaly,nObs=nObs,tx=tx,

```



```

        nTx=nTx,study=study,se1=se1,se2=se2,EQ-5D=EQ-
5D,gender=gender,nStudies=nStudies,mixed=0)

AggBugsData <- list(cost=cost,qaly=qaly,nObs=nObs,tx=tx,
        nTx=nTx,se1=se1,se2=se2,EQ-5D=EQ-
5D,gender=gender,study=study,nStudies=nStudies,

aggNobs=aggNobs,aggSe2=aggSe2,aggSeSD=aggSeSD,aggN=aggN,mixed=0)

parameters.to.save <-
c("deltaCost","deltaQaly","ctBeta","qtBeta","csBeta","qsBeta","s2tBeta","cSe","qSe","cAlpha"
)

#parameters.to.save <- c("deltaCost","deltaQaly")

#model 1 Direct
WIN.SOURCE <- 'GR7NSHM1.txt'
res1 <- genSamps(data=bugsData,init=init,parameters.to.save=parameters.to.save)

#model 2 Indirect

WIN.SOURCE <- 'GR7NSHM2.txt'
bugsData$mixed<-0
res2 <- genSamps(data=bugsData,init=init,parameters.to.save=parameters.to.save)

#model 3 Mixed
WIN.SOURCE <- 'GR7NSHM2.txt'
bugsData$mixed<-1
res3 <- genSamps(data=bugsData,init=init,parameters.to.save=parameters.to.save)

#model 2 Indirect + Agg
WIN.SOURCE <- 'GR7NSHM2Agg.txt'
AggBugsData$mixed<-0
res2Agg <-
genSamps(data=AggBugsData,init=init,parameters.to.save=parameters.to.save)

#model 3 Mixed + Agg
WIN.SOURCE <- 'GR7NSHM2Agg.txt'
AggBugsData$mixed<-1
res3Agg <-
genSamps(data=AggBugsData,init=init,parameters.to.save=parameters.to.save)

#save.image("C:\\work\\phd_stuff\\new.RData")
save.image("C:\\work\\phd_stuff\\finished models oct 2007\\new.RData")

ncost1 <- res1$sims.matrix[, "deltaCost"]
ncost2 <- res2$sims.matrix[, "deltaCost"]
ncost3 <- res3$sims.matrix[, "deltaCost"]
ncost2agg <- res2Agg$sims.matrix[, "deltaCost"]
ncost3agg <- res3Agg$sims.matrix[, "deltaCost"]

```

```

qaly1 <- res1$sims.matrix[, "deltaQaly"]
qaly2 <- res2$sims.matrix[, "deltaQaly"]
qaly3 <- res3$sims.matrix[, "deltaQaly"]
qaly2agg <- res2Agg$sims.matrix[, "deltaQaly"]
qaly3agg <- res3Agg$sims.matrix[, "deltaQaly"]

```

```

ctBeta1 <- res1$sims.matrix[, "ctBeta"]
ctBeta2 <- res2$sims.matrix[, "ctBeta"]
ctBeta3 <- res3$sims.matrix[, "ctBeta"]
ctBeta2agg <- res2Agg$sims.matrix[, "ctBeta"]
ctBeta3agg <- res3Agg$sims.matrix[, "ctBeta"]

```

```

qtBeta1 <- res1$sims.matrix[, "qtBeta"]
qtBeta2 <- res2$sims.matrix[, "qtBeta"]
qtBeta3 <- res3$sims.matrix[, "qtBeta"]
qtBeta2agg <- res2Agg$sims.matrix[, "qtBeta"]
qtBeta3agg <- res3Agg$sims.matrix[, "qtBeta"]

```

```

csBeta1 <- res1$sims.matrix[, "csBeta"]
csBeta2 <- res2$sims.matrix[, "csBeta"]
csBeta3 <- res3$sims.matrix[, "csBeta"]
csBeta2agg <- res2Agg$sims.matrix[, "csBeta"]
csBeta3agg <- res3Agg$sims.matrix[, "csBeta"]

```

```

qsBeta1 <- res1$sims.matrix[, "qsBeta"]
qsBeta2 <- res2$sims.matrix[, "qsBeta"]
qsBeta3 <- res3$sims.matrix[, "qsBeta"]
qsBeta2agg <- res2Agg$sims.matrix[, "qsBeta"]
qsBeta3agg <- res3Agg$sims.matrix[, "qsBeta"]

```

```

stBeta1 <- res1$sims.matrix[, "s2tBeta"]
stBeta2 <- res2$sims.matrix[, "s2tBeta"]
stBeta3 <- res3$sims.matrix[, "s2tBeta"]
stBeta2agg <- res2Agg$sims.matrix[, "s2tBeta"]
stBeta3agg <- res3Agg$sims.matrix[, "s2tBeta"]

```

```

#intCost <- rnorm(N.SIMS, 250, 50)
intCost <- 198

```

```

cost1 <- ncost1+intCost
cost2 <- ncost2+intCost
cost3 <- ncost3+intCost
cost2agg <- ncost2agg+intCost
cost3agg <- ncost3agg+intCost

```

```

ceRange <- seq(0, 100000, 100)
ceac1 <- numeric()
ceac2 <- numeric()
ceac3 <- numeric()
ceac2agg <- numeric()
ceac3agg <- numeric()

```

```

icer<-numeric()

```

```

evpi1 <- numeric()

```



```

evpi2 <- numeric()
evpi3 <- numeric()
evpi2agg <- numeric()
evpi3agg <- numeric()

for (cc in seq(along=ceRange)){

  lambda <- ceRange[cc]

  nb1<-qaly1*lambda-cost1
  nb2<-qaly2*lambda-cost2
  nb3<-qaly3*lambda-cost3
  nb2agg<-qaly2agg*lambda-cost2agg
  nb3agg<-qaly3agg*lambda-cost3agg

  ceac1[cc] <- mean(nb1>0)
  ceac2[cc] <- mean(nb2>0)
  ceac3[cc] <- mean(nb3>0)
  ceac2agg[cc] <- mean(nb2agg>0)
  ceac3agg[cc] <- mean(nb3agg>0)

  evpi1[cc] <- sum(nb1[nb1>0]/N.SIMS)-max(0,sum(nb1/N.SIMS))
  evpi2[cc] <- sum(nb2[nb2>0]/N.SIMS)-max(0,sum(nb2/N.SIMS))
  evpi3[cc] <- sum(nb3[nb3>0]/N.SIMS)-max(0,sum(nb3/N.SIMS))
  evpi2agg[cc] <- sum(nb2agg[nb2agg>0]/N.SIMS)-max(0,sum(nb2agg/N.SIMS))
  evpi3agg[cc] <- sum(nb3agg[nb3agg>0]/N.SIMS)-max(0,sum(nb3agg/N.SIMS))

  #nb1[nb1<0]<-0
  #nb2[nb2<0]<-0

  #modelVoi[cc,ii] <- mean(nb)-mean(seNb)
}

plot(0,0,xlim=range(ceRange)/1000,ylim=c(0,1),xlab="Cost-effectiveness Threshold
(£1000,s)",ylab="Probability EPP
cost-effective",bty="l",type="n")
lines(ceRange/1000,ceac1,col=1,lwd=6,lty=1)
lines(ceRange/1000,ceac2,col=2,lwd=6,lty=2)
lines(ceRange/1000,ceac3,col=3,lwd=6,lty=3)
legend(locator(1),legend=c("Direct","Joint"),lty=c(1,3), bty="n",lwd=4,col=c(1,3),cex=1.2)
text(locator(1),"Indirect",cex=1.2)

par(cex=1.2)
plot(0,0,xlim=range(ceRange/1000),ylim=c(0,1),xlab="Cost-effectiveness Threshold
(£1000,s)",ylab="Probability EPP cost-effective",bty="l",type="n")

lines(ceRange/1000,ceac2,col=2,lwd=6,lty=1)
lines(ceRange/1000,ceac2agg,col=2,lwd=6,lty=2)
lines(ceRange/1000,ceac3,col=4,lwd=6,lty=1)
lines(ceRange/1000,ceac3agg,col=4,lwd=6,lty=2)
lines(ceRange/1000,ceac1,col=1,lwd=6,lty=1)

text(locator(1),"Joint + Add",cex=1.2,pos=2)

```

```

text(locator(1),"Joint",cex=1.2,pos=4)
text(locator(1),"Direct",cex=1.2,pos=4)

text(locator(1),"Indirect + Add",cex=1.2,pos=4)
text(locator(1),"Indirect",cex=1.2,pos=4)

plot(0,0,xlim=range(ceRange/1000),ylim=c(0,150),xlab="Cost-Effectiveness Threshold
(£1000,s)",ylab="EVPI",bty="l",type="n")
lines(ceRange/1000,evpi1,col=1,lwd=6,lty=1)
lines(ceRange/1000,evpi2,col=2,lwd=6,lty=1)
lines(ceRange/1000,evpi2agg,col=2,lwd=6,lty=3)
lines(ceRange/1000,evpi3,col=4,lwd=6,lty=1)
lines(ceRange/1000,evpi3agg,col=4,lwd=6,lty=3)
# legend(locator(1),legend=c("Direct","Indirect","Indirect + Agg","Joint","Joint + Agg"),
# bty="n",lwd=4,col=c(1,2,2,3,3),lty=c(1,1,3,1,3),cex=1.2)

text(locator(1),"Joint + Add",cex=1.2,pos=2)
text(locator(1),"Joint",cex=1.2,pos=4)
text(locator(1),"Direct",cex=1.2,pos=4)

text(locator(1),"Indirect + Add",cex=1.2,pos=2)
text(locator(1),"Indirect",cex=1.2,pos=4)

results <- rbind(
  c(mean(cost1),quantile(cost1,c(0.025,0.975))),
  mean(cost2),quantile(cost2,c(0.025,0.975)),
  mean(cost3),quantile(cost3,c(0.025,0.975)),
  mean(cost2agg),quantile(cost2agg,c(0.025,0.975)),
  mean(cost3agg),quantile(cost3agg,c(0.025,0.975))),
  c(mean(qaly1),quantile(qaly1,c(0.025,0.975))),
  mean(qaly2),quantile(qaly2,c(0.025,0.975)),
  mean(qaly3),quantile(qaly3,c(0.025,0.975)),
  mean(qaly2agg),quantile(qaly2agg,c(0.025,0.975)),
  mean(qaly3agg),quantile(qaly3agg,c(0.025,0.975))),
  c(mean(ctBeta1),quantile(ctBeta1,c(0.025,0.975))),
  mean(ctBeta2),quantile(ctBeta2,c(0.025,0.975)),
  mean(ctBeta3),quantile(ctBeta3,c(0.025,0.975)),
  mean(ctBeta2agg),quantile(ctBeta2agg,c(0.025,0.975)),
  mean(ctBeta3agg),quantile(ctBeta3agg,c(0.025,0.975))),
  c(mean(qtBeta1),quantile(qtBeta1,c(0.025,0.975))),
  mean(qtBeta2),quantile(qtBeta2,c(0.025,0.975)),
  mean(qtBeta3),quantile(qtBeta3,c(0.025,0.975)),
  mean(qtBeta2agg),quantile(qtBeta2agg,c(0.025,0.975)),
  mean(qtBeta3agg),quantile(qtBeta3agg,c(0.025,0.975))),
  c(mean(csBeta1),quantile(csBeta1,c(0.025,0.975))),
  mean(csBeta2),quantile(csBeta2,c(0.025,0.975)),
  mean(csBeta3),quantile(csBeta3,c(0.025,0.975)),
  mean(csBeta2agg),quantile(csBeta2agg,c(0.025,0.975)),
  mean(csBeta3agg),quantile(csBeta3agg,c(0.025,0.975))),
  c(mean(qsBeta1),quantile(qsBeta1,c(0.025,0.975))),
  mean(qsBeta2),quantile(qsBeta2,c(0.025,0.975)),
  mean(qsBeta3),quantile(qsBeta3,c(0.025,0.975)),

```



```

mean(qsBeta2agg),quantile(qsBeta2agg,c(0.025,0.975)),
mean(qsBeta3agg),quantile(qsBeta3agg,c(0.025,0.975))),
c(mean(stBeta1),quantile(stBeta1,c(0.025,0.975)),
mean(stBeta2),quantile(stBeta2,c(0.025,0.975)),
mean(stBeta3),quantile(stBeta3,c(0.025,0.975)),
mean(stBeta2agg),quantile(stBeta2agg,c(0.025,0.975)),
mean(stBeta3agg),quantile(stBeta3agg,c(0.025,0.975))))

results[1,]<-round(results[1,],0)
results[2,]<-round(results[2,],3)
results[3,]<-round(results[3,],0)
results[4,]<-round(results[4,],3)
results[5,]<-round(results[5,],0)
results[6,]<-round(results[6,],3)
results <- rbind(results,results[1,]/results[2,])

colnames(results) <-
c("Direct","2.5%","97.5%","Indirect","2.5%","97.5%","Joint","2.5%","97.5%","Indirect+Agg","2
.5%","97.5%","Joint+Agg","2.5%","97.5%")
rownames(results) <- c("Cost","QALY","BCT","BQT","BCS","BQS","BST","ICER")

layout(1)

par(new=FALSE)
plot(qaly1,cost1,ylab="Inc. Cost",xlab="Inc. Qaly",ylim=c(0,400),xlim=c(-
0.02,0.02),bty="n",main="Direct Analysis")
par(new=TRUE)
curve(30000*x,ylim=c(0,400),xlim=c(-0.02,0.02),ylab=" ",xlab=" ",bty="n",lwd=3,col="grey")
abline(h=0,lwd=1)
abline(v=0,lwd=1)

par(new=FALSE)
plot(qaly2,cost2,ylab="Inc. Cost",xlab="Inc. Qaly",ylim=c(0,400),xlim=c(-
0.02,0.02),bty="n",main="Indirect Analysis")
par(new=TRUE)
curve(30000*x,ylim=c(0,400),xlim=c(-0.02,0.02),ylab=" ",xlab=" ",bty="n",lwd=3,col="grey")
abline(h=0,lwd=1)
abline(v=0,lwd=1)

par(new=FALSE)
plot(qaly3,cost3,ylab="Inc. Cost",xlab="Inc. Qaly",ylim=c(0,400),xlim=c(-
0.02,0.02),bty="n",main="Joint Analysis")
par(new=TRUE)
curve(30000*x,ylim=c(0,400),xlim=c(-0.02,0.02),ylab=" ",xlab=" ",bty="n",lwd=3,col="grey")
abline(h=0,lwd=1)
abline(v=0,lwd=1)

par(new=FALSE)
plot(qaly3agg,cost3agg,ylab="Inc. Cost",xlab="Inc. Qaly",ylim=c(0,400),xlim=c(-
0.02,0.02),bty="n",main="Joint Analysis")
par(new=TRUE)
curve(30000*x,ylim=c(0,400),xlim=c(-0.02,0.02),ylab=" ",xlab=" ",bty="n",lwd=3,col="grey")
abline(h=0,lwd=1)

```

```
abline(v=0,lwd=1)
```

E.2 WinBUGS code for evidence synthesis in Chapter 6

E.2.1 Direct Model

```
Model{
  for(ii in 1:nObs){
    #MODEL 1: TRIAL BASED ANALYSIS

    cMean[ii] <- cAlpha[study[ii]] + ctBeta*tx[ii] + ceBeta[study[ii]]*EQ-5D[ii] +
    corrBeta[study[ii],tx[ii]+1] * qaly[ii]

    logCMean[ii] <- log(cMean[ii])-0.5*cVar[study[ii]]
    cost[ii] ~ dlnorm(logCMean[ii],cPrec[study[ii]]) #trial c mean and precision

    qaly[ii] ~ dnorm(qMean[ii], qPrec[study[ii]]) #trial q mean,precision
    qMean[ii] <- qAlpha[study[ii]] + qtBeta*tx[ii] + qeBeta[study[ii]]*EQ-5D[ii]

    se1[ii]~dnorm(0,0.01)
    se2[ii]~dnorm(0,0.01)

    gender[ii] ~ dnorm(0,0.01)
  }

  dummy1 <- mixed

  csBeta ~dunif(0,0)
  qsBeta ~dunif(0,0)
  s2tBeta ~dunif(1,1)

  for (ss in 1:nStudies)      {
    for (tt in 1:nTx)      {
      corrBeta[ss,tt] ~ dnorm (0,1.0E-3)
      #      psi[ss,tt] <- corrBeta[ss,tt]*sqrt(qvar[ss])/sqrt(cvar[ss])
    }

    ceBeta[ss]~ dnorm (0,1.0E-12)
    qeBeta[ss]~ dnorm (0,1.0E-12)

    cAlpha[ss] ~ dnorm (0,1.0E-12)
    qAlpha[ss] ~ dnorm (0,1.0E-12)
    seAlpha[ss] ~ dnorm (0,1.0E-12)

    s1Alpha[ss]~ dnorm (0,1.0E-12)
    s2Alpha[ss]~ dnorm (0,1.0E-12)

    s1pSe[ss] ~ dunif (0,10)
    s2pSe[ss] ~ dunif (0,10)

    qSe[ss] ~ dunif (0,5)
  }
}
```



```

        qVar[ss] <- pow(qSe[ss],2)
        qPrec[ss] <- 1/qVar[ss]

        logCostSe[ss] ~ dunif (-5,10)
        cSe[ss] <- exp(logCostSe[ss])
        cVar[ss] <- pow(cSe[ss],2)
        cPrec[ss]<-1/cVar[ss]

        sepSe[ss] ~ dunif (0,10)
        sePrec[ss]<-pow(sepSe[ss], -2)
    }

    deltaCost <- ctBeta
    deltaQaly <- qtBeta

    ctBeta~ dnorm (0,1.0E-12)
    qtBeta~ dnorm (0,1.0E-12)
}

```

E.2.2 Indirect and Joint models

```

Model{
  for(ii in 1:nObs){
    #MODEL 1: TRIAL BASED ANALYSIS

    cMean[ii] <- cAlpha[study[ii]] + ctBeta*tx[ii] *mixed+ corrBeta[study[ii],tx[ii]+1] *
qaly[ii] + ceBeta[study[ii]]*EQ-5D[ii]+ csBeta*se2[ii]

    logCMean[ii] <- log(cMean[ii])-0.5*cVar[study[ii]]
    cost[ii] ~ dlnorm(logCMean[ii],cPrec[study[ii]]) #trial c mean and precision

    qaly[ii] ~ dnorm(qMean[ii], qPrec[study[ii]]) #trial q mean,precision
    qMean[ii] <- qAlpha[study[ii]] + qtBeta*tx[ii] *mixed+ qsBeta*se2[ii] +
qeBeta[study[ii]]*EQ-5D[ii]

    #SELF-EFFICACY IN TERMS OF TREATMENT GROUP
    #SELF-EFFICACY IN TERMS OF TREATMENT GROUP
    se1[ii]~dnorm(s1Alpha[study[ii]],s1Prec[study[ii]])

    gender[ii] ~ dnorm(0,0.01)

    se2[ii]~dnorm(s2Mean[ii],s2Prec[study[ii]])
    s2Mean[ii]<-s2Alpha[study[ii]] +s2tBeta*tx[ii]

  }

  for (aa in 1:aggNobs){
    aggSe2[aa] ~ dnorm(s2tBeta,aggSeTau[aa])
    aggSeTau[aa] <- aggN[aa]/pow(aggSeSD[aa],2)
  }
}

```

```

for (ss in 1:nStudies)      {
  for (tt in 1:nTx)        {
    corrBeta[ss,tt] ~ dnorm (0,1.0E-3)
    #   psi[ss,tt] <- corrBeta[ss,tt]*sqrt(qvar[ss])/sqrt(cvar[ss])
  }

  qeBeta[ss]~ dnorm (0,1.0E-12)
  ceBeta[ss]~ dnorm (0,1.0E-12)

  s1Alpha[ss]~ dnorm (0,1.0E-12)

  s2Alpha[ss]~ dnorm (0,1.0E-12)

  cAlpha[ss] ~ dnorm (0,1.0E-12)
  qAlpha[ss] ~ dnorm (0,1.0E-12)

  qSe[ss] ~ dunif (0,5)
  qVar[ss] <- pow(qSe[ss],2)
  qPrec[ss] <- 1/qVar[ss]

  sSe[ss] ~ dunif (0,5)
  sVar[ss] <- pow(sSe[ss],2)
  sPrec[ss] <- 1/sVar[ss]

  logCostSe[ss] ~ dunif (-5,10)
  cSe[ss] <- exp(logCostSe[ss])
  cVar[ss] <- pow(cSe[ss],2)
  cPrec[ss]<-1/cVar[ss]

  s1pSe[ss] ~ dunif (0,10)
  s1Prec[ss]<-pow(s1pSe[ss], -2)

  s2pSe[ss] ~ dunif (0,10)
  s2Prec[ss]<-pow(s2pSe[ss], -2)

}

deltaCost <- ctBeta*mixed+csBeta*s2tBeta
deltaQaly <- qtBeta*mixed+qsBeta*s2tBeta

ctBeta~ dnorm (0,1.0E-12)
qtBeta~ dnorm (0,1.0E-12)

csBeta~ dnorm (0,1.0E-12)
qsBeta~ dnorm (0,1.0E-12)

s2tBeta~ dnorm (0,1.0E-12)

}

```


Appendix F. Discrete Choice Experiment Questionnaires

How to fill in the questionnaire

In this questionnaire we ask you to make choices between the pairs of options (Option A and Option B) that are presented to you. Each option has four characteristics associated with it. These four characteristics are:

- your health state
- how confident you are in your ability to manage your condition
- how long you have to wait for a GP appointment
- your contact with friends and relatives

We ask you to consider ALL these characteristics when making your decision.

IMPORTANT

For each question please choose the **option you would prefer, NOT the option you feel best describes your current situation** (i.e. please tick either A or B). There are no right and wrong answers.

Please treat each question separately and **imagine** that these are choices that are actually available to you.

Patient identification number

--	--	--	--

Here is an example:

Imagine that you can have either option A or option B, which would you choose? If you would prefer the option where you have **moderate** pain or discomfort, are **not** confident that you can manage your condition but you can have a GP appointment **tomorrow** and you see friends or relatives **daily** (i.e. everything in box A), then choose option A.

However, if you would prefer the option where you have **moderate** pain or discomfort as well as having **some** problems with walking, but you are **totally** confident that you can manage your condition with a GP appointment in **one week's time**, but you **rarely** see friends or relatives (i.e. everything in box B), then choose option B.

- A**
- You have
no problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
 - You are **not** confident you can manage your condition
 - You can have a GP appointment **tomorrow**
 - You see your friends or relatives **daily**

OR

- B**
- You have
some problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
 - You are **totally** confident you can manage your condition
 - You can have a GP appointment in **one week's time**
 - You **rarely** see friends or relatives

Please tick one box:

Choice A

☐

Choice B

☐

Now please turn to the next page to start filling in the questionnaire.

Please choose the option (A or B) that you would prefer. Remember to imagine that these options are actually available to you.

Question 1: If you had to choose between A and B below, which would you choose?

A

- You have
no problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **moderately** confident you can manage your condition
- You have a GP appointment in **one week's** time
- You **rarely** see friends or relatives.

OR

B

- You have
some problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **not** confident you can manage your condition
- You have a GP appointment in **three week's** time
- You see friends or relatives **daily**

Please tick one box:

Choice A

☐

Choice B

☐

Question 2: If you had to choose between A and B below, which would you choose?

A

- You have
no problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **not** confident you can manage your condition
- You have a GP appointment in **one week's** time
- You see friends or relatives every **few days**

OR

B

- You have
some problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **totally** confident you can manage your condition
- You have a GP appointment **tomorrow**
- You **rarely** see friends or relatives

Please tick one box:

Choice A

☐

Choice B

☐

Please choose the option (A or B) that you would prefer. Remember to imagine that these options are actually available to you.

Question 3: If you had to choose between A and B below, which would you choose?

A

- You have **some** problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **totally** confident you can manage your condition
- You have a GP appointment in **one week's** time
- You see friends or relatives every **few days**

OR

B

- You have **some** problems walking about
some problems with self care
no problems with usual activities
moderate pain or discomfort
moderate anxiety or depression
- You are **moderately** confident you can manage your condition
- You have a GP appointment in **three week's** time
- You **rarely** see friends or relatives.

Please tick one box:

Choice A

☐

Choice B

☐

Question 4: If you had to choose between A and B below, which would you choose?

A

- You have **some** problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **moderately** confident you can manage your condition
- You have a GP appointment in **three week's** time
- You see friends or relatives **daily**

OR

B

- You have **some** problems walking about
some problems with self care
no problems with usual activities
moderate pain or discomfort
moderate anxiety or depression
- You are **not** confident you can manage your condition
- You have a GP appointment **tomorrow**
- You see friends or relatives every **few days**

Please tick one box:

Choice A

☐

Choice B

☐

Please choose the option (A or B) that you would prefer. Remember to imagine that these options are actually available to you.

Question 5: If you had to choose between A and B below, which would you choose?

A

- You have
no problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **totally** confident you can
manage your condition
- You have a GP appointment
tomorrow
- You see friends or relatives **daily**

B

- You have
some problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **moderately** confident you
can manage your condition
- You have a GP appointment in
one week's time
- You see friends or relatives every
few days

OR

Please tick one box:

Choice A

☐

Choice B

☐

Question 6: If you had to choose between A and B below, which would you choose?

A

- You have
some problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **not** confident you can
manage your condition
- You have a GP appointment
tomorrow
- You **rarely** see friends or relatives

B

- You have
some problems walking about
some problems with self care
no problems with usual activities
moderate pain or discomfort
moderate anxiety or depression
- You are **totally** confident you can
manage your condition
- You have a GP appointment in **one**
week's time
- You see friends or relatives **daily**

OR

Please tick one box:

Choice A

☐

Choice B

☐

Question 7: If you had to choose between A and B below, which would you choose?

A

- You have **some** problems walking about
some problems with self care
no problems with usual activities
moderate pain or discomfort
moderate anxiety or depression
- You are **totally** confident you can manage your condition
- You have a GP appointment in **three week's** time
- You **rarely** see friends or relatives

OR

B

- You have no problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **moderately** confident you can manage your condition
- You have a GP appointment **tomorrow**
- You see friends or relatives **daily**

Please tick one box:

Choice A

☐

Choice B

☐

Question 8: If you had to choose between A and B below, which would you choose?

A

- You have **some** problems walking about
some problems with self care
no problems with usual activities
moderate pain or discomfort
moderate anxiety or depression
- You are **moderately** confident you can manage your condition
- You have a GP appointment **tomorrow**
- You see friends or relatives every **few days**

OR

B

- You have no problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **not** confident you can manage your condition
- You have a GP appointment in **one week's** time
- You **rarely** see friends or relatives

Please tick one box:

Choice A

☐

Choice B

☐

MANY THANKS FOR TAKING THE TIME TO COMPLETE THIS FORM

Please choose the option (A or B) that you would prefer. Remember to imagine that these options are actually available to you.

Question 9: If you had to choose between A and B below, which would you choose?

A

- You have **some** problems walking about
some problems with self care
no problems with usual activities
moderate pain or discomfort
moderate anxiety or depression
- You are **not** confident you can manage your condition
- You have a GP appointment in **one week's** time
- You see friends or relatives **daily**

OR

B

- You have
no problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **totally** confident you can manage your condition
- You have a GP appointment in **three week's** time
- You see friends or relatives every **few days**

Please tick one box:

Choice A

☐

Choice B

☐

Question 10: If you had to choose between A and B below, which would you choose?

A

- You have
some problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **totally** confident you can manage your condition
- You have a GP appointment in **one week's** time
- You see friends or relatives every **few days**

OR

B

- You have
some problems walking about
some problems with self care
no problems with usual activities
moderate pain or discomfort
moderate anxiety or depression
- You are **moderately** confident you can manage your condition
- You have a GP appointment in **three week's** time
- You **rarely** see friends or relatives

Please tick one box:

Choice A

☐

Choice B

☐

MANY THANKS FOR TAKING THE TIME TO COMPLETE THIS FORM



How to fill in the questionnaire

In this questionnaire we ask you to make choices between the pairs of options (Option A and Option B) that are presented to you. Each option has four characteristics associated with it. These four characteristics are:

- your health state
- how confident you are in your ability to manage your condition
- how long you have to wait for a GP appointment
- your contact with friends and relatives

We ask you to consider ALL these characteristics when making your decision.

IMPORTANT

For each question please choose the **option you would prefer**, **NOT the option you feel best describes your current situation** (i.e. please tick either A or B). There are no right and wrong answers.

Please treat each question separately and **imagine** that these are choices that are actually available to you.

Now please turn to the next page to start filling in the questionnaire.

Patient identification number

--	--	--	--



Here is an example:

Imagine that you can have either option A or option B, which would you choose? If you would prefer the option where you have **moderate** pain or discomfort, are **not** confident that you can manage your condition but you can have a GP appointment **tomorrow** and you see friends or relatives **daily** (i.e. everything in box A), then choose option A.

However, if you would prefer the option where you have **moderate** pain or discomfort as well as having **some** problems with walking, but you are **totally** confident that you can manage your condition with a GP appointment in **one week's time**, but you **rarely** see friends or relatives (i.e. everything in box B), then choose option B.

A

- You have
no problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **not** confident you can manage your condition
- You can have a GP appointment **tomorrow**
- You see your friends or relatives **daily**

OR

B

- You have
some problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **totally** confident you can manage your condition
- You can have a GP appointment in **one week's time**
- You **rarely** see friends or relatives

Please tick one box:

Choice A

☐

Choice B

☐

Now please turn to the next page to start filling in the questionnaire.

Question 1: If you had to choose between A and B below, which would you choose?

A

- You have
no problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **totally** confident you can
manage your condition
- You can have a GP appointment in
one week's time
- You **rarely** see your friends or
relatives

OR

B

- You have
some problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **moderately** confident you
can manage your condition
- You can have a GP appointment in
three week's time
- You see friends or relatives **daily**

Please tick one box:

Choice A

☐

Choice B

☐

Question 2: If you had to choose between A and B below, which would you choose?

A

- You have
no problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **totally** confident you can
manage your condition
- You can have a GP appointment in
one week's time
- You see friends or relatives every
few days

OR

B

- You have
some problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **moderately** confident you
can manage your condition
- You can have a GP appointment
tomorrow
- You **rarely** see your friends or
relatives

Please tick one box:

Choice A

☐

Choice B

☐

Question 3: If you had to choose between A and B below, which would you choose?

A

- You have
no problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **moderately** confident you
can manage your condition
- You can have a GP appointment in
tomorrow
- You **rarely** see your friends or
relatives

OR

B

- You have
some problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **not** confident you can
manage your condition
- You can have a GP appointment in
one week's time
- You see friends or relatives **daily**

Please tick one box:

Choice A

☐

Choice B

☐

Question 4: If you had to choose between A and B below, which would you choose?

A

- You have
no problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **moderately** confident you
can manage your condition
- You can have a GP appointment in
one week's time
- You see friends or relatives every
few days

OR

B

- You have
some problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **not** confident you can
manage your condition
- You can have a GP appointment in
three week's time
- You **rarely** see your friends or
relatives

Please tick one box:

Choice A

☐

Choice B

☐

Question 5: If you had to choose between A and B below, which would you choose?

A

- You have
no problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **moderately** confident you can manage your condition
- You can have a GP appointment in **three week's** time
- You see friends or relatives **daily**

OR

B

- You have
some problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **not** confident you can manage your condition
- You can have a GP appointment **tomorrow**
- You see friends or relatives every **few days**

Please tick one box:

Choice A

☐

Choice B

☐

Question 6: If you had to choose between A and B below, which would you choose?

A

- You have
no problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **not** confident you can manage your condition
- You can have a GP appointment **tomorrow**
- You see friends or relatives every **few days**

OR

B

- You have
some problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **totally** confident you can manage your condition
- You can have a GP appointment in **one week's** time
- You **rarely** see your friends or relatives

Please tick one box:

Choice A

☐

Choice B

☐

Question 7: If you had to choose between A and B below, which would you choose?

A

- You have
no problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **not** confident you can manage your condition
- You can have a GP appointment in **one week's** time
- You see friends or relatives **daily**
few days

OR

B

- You have
some problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **totally** confident you can manage your condition
- You can have a GP appointment in **three week's** time
- You see friends or relatives every **few days**

Please tick one box:

Choice A

☐

Choice B

☐

Question 8: If you had to choose between A and B below, which would you choose?

A

- You have
no problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **not** confident you can manage your condition
- You can have a GP appointment in **three week's** time
- You **rarely** see your friends or relatives

OR

B

- You have
some problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **totally** confident you can manage your condition
- You can have a GP appointment **tomorrow**
- You see friends or relatives **daily**

Please tick one box:

Choice A

☐

Choice B

☐

Question 9: If you had to choose between A and B below, which would you choose?

A

- You have **some** problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **totally** confident you can manage your condition
- You can have a GP appointment **tomorrow**
- You see friends or relatives every **few days**

OR

B

- You have **some** problems walking about
some problems with self care
no problems with usual activities
moderate pain or discomfort
moderate anxiety or depression
- You are **moderately** confident you can manage your condition
- You can have a GP appointment in **one week's** time
- You **rarely** see your friends or relatives

Please tick one box:

Choice A

☐

Choice B

☐

Question 10: If you had to choose between A and B below, which would you choose?

A

- You have **some** problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **totally** confident you can manage your condition
- You can have a GP appointment in **one week's** time
- You see friends or relatives **daily**

OR

B

- You have **some** problems walking about
some problems with self care
no problems with usual activities
moderate pain or discomfort
moderate anxiety or depression
- You are **moderately** confident you can manage your condition
- You can have a GP appointment in **three week's** time
- You see friends or relatives every **few days**

Please tick one box:

Choice A

☐

Choice B

☐

Question 11: If you had to choose between A and B below, which would you choose?

A

- You have **some** problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **totally** confident you can manage your condition
- You can have a GP appointment in **three week's** time
- You **rarely** see your friends or relatives

OR

B

- You have **some** problems walking about
some problems with self care
no problems with usual activities
moderate pain or discomfort
moderate anxiety or depression
- You are **moderately** confident you can manage your condition
- You can have a GP appointment **tomorrow**
- You see friends or relatives **daily**

Please tick one box:

Choice A ☐

Choice B ☐

Question 12: If you had to choose between A and B below, which would you choose?

A

- You have
no problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **totally** confident you can manage your condition
- You can have a GP appointment **tomorrow**
- You see friends or relatives **daily**

OR

B

- You have
some problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **moderately** confident you can manage your condition
- You can have a GP appointment in **one week's** time
- You see friends or relatives every **few days**

Please tick one box:

Choice A ☐

Choice B ☐

Question 13: If you had to choose between A and B below, which would you choose?

- A**
- You have **some** problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
 - You are **moderately** confident you can manage your condition
 - You can have a GP appointment **tomorrow**
 - You see friends or relatives **daily**

OR

- B**
- You have **some** problems walking about
some problems with self care
no problems with usual activities
moderate pain or discomfort
moderate anxiety or depression
 - You are **not** confident you can manage your condition
 - You can have a GP appointment in **one week's time**
 - You see friends or relatives every **few days**

Please tick one box:

Choice A ☐

Choice B ☐

Question 14: If you had to choose between A and B below, which would you choose?

- A**
- You have **some** problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
 - You are **moderately** confident you can manage your condition
 - You can have a GP appointment in **one week's time**
 - You **rarely** see your friends or relatives

OR

- B**
- You have **some** problems walking about
some problems with self care
no problems with usual activities
moderate pain or discomfort
moderate anxiety or depression
 - You are **not** confident you can manage your condition
 - You can have a GP appointment in **three week's time**
 - You see friends or relatives **daily**

Please tick one box:

Choice A ☐

Choice B ☐

Question 15: If you had to choose between A and B below, which would you choose?

A

- You have
some problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **moderately** confident you can manage your condition
- You can have a GP appointment in **three week's** time
- You see friends or relatives every **few days**

OR

B

- You have
some problems walking about
some problems with self care
no problems with usual activities
moderate pain or discomfort
moderate anxiety or depression
- You are **not** confident you can manage your condition
- You can have a GP appointment **tomorrow**
- You **rarely** see your friends or relatives

Please tick one box:

Choice A

☐

Choice B

☐

Question 16: If you had to choose between A and B below, which would you choose?

A

- You have
some problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **not** confident you can manage your condition
- You can have a GP appointment **tomorrow**
- You **rarely** see your friends or relatives

OR

B

- You have
some problems walking about
some problems with self care
no problems with usual activities
moderate pain or discomfort
moderate anxiety or depression
- You are **totally** confident you can manage your condition
- You can have a GP appointment in **one week's** time
- You see friends or relatives **daily**

Please tick one box:

Choice A

☐

Choice B

☐

Question 17: If you had to choose between A and B below, which would you choose?

A

- You have **some** problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **not** confident you can manage your condition
- You can have a GP appointment in **one week's** time
- You see friends or relatives every **few days**

OR

B

- You have **some** problems walking about
some problems with self care
no problems with usual activities
moderate pain or discomfort
moderate anxiety or depression
- You are **totally** confident you can manage your condition
- You can have a GP appointment in **three week's** time
- You **rarely** see your friends or relatives

Please tick one box:

Choice A

☐

Choice B

☐

Question 18: If you had to choose between A and B below, which would you choose?

A

- You have **some** problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **not** confident you can manage your condition
- You can have a GP appointment in **three week's** time
- You see friends or relatives **daily**

OR

B

- You have **some** problems walking about
some problems with self care
no problems with usual activities
moderate pain or discomfort
moderate anxiety or depression
- You are **totally** confident you can manage your condition
- You can have a GP appointment **tomorrow**
- You see friends or relatives every **few days**

Please tick one box:

Choice A

☐

Choice B

☐

Question 19: If you had to choose between A and B below, which would you choose?

A

- You have **some** problems walking about
some problems with self care
no problems with usual activities
moderate pain or discomfort
moderate anxiety or depression
- You are **totally** confident you can manage your condition
- You can have a GP appointment **tomorrow**
- You **rarely** see your friends or relatives

OR

B

- You have
no problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **moderately** confident you can manage your condition
- You can have a GP appointment in **one week's** time
- You see friends or relatives **daily**

Please tick one box:

Choice A

☐

Choice B

☐

Question 20: If you had to choose between A and B below, which would you choose?

A

- You have
some problems walking about
some problems with self care
no problems with usual activities
moderate pain or discomfort
moderate anxiety or depression
- You are **totally** confident you can manage your condition
- You can have a GP appointment in **one week's** time
- You see friends or relatives every **few days**

OR

B

- You have
no problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **moderately** confident you can manage your condition
- You can have a GP appointment in **three week's** time
- You **rarely** see your friends or relatives

Please tick one box:

Choice A

☐

Choice B

☐

Question 21: If you had to choose between A and B below, which would you choose?

- A**
- You have **some** problems walking about **some** problems with self care no problems with usual activities **moderate** pain or discomfort **moderate** anxiety or depression
 - You are **totally** confident you can manage your condition
 - You can have a GP appointment in **three week's** time
 - You see friends or relatives **daily**

OR

- B**
- You have no problems walking about no problems with self care no problems with usual activities **moderate** pain or discomfort no anxiety or depression
 - You are **moderately** confident you can manage your condition
 - You can have a GP appointment **tomorrow**
 - You see friends or relatives every **few days**

Please tick one box:

Choice A

☐

Choice B

☐

Question 22: If you had to choose between A and B below, which would you choose?

- A**
- You have **some** problems walking about **some** problems with self care no problems with usual activities **moderate** pain or discomfort **moderate** anxiety or depression
 - You are **moderately** confident you can manage your condition
 - You can have a GP appointment **tomorrow**
 - You see friends or relatives every **few days**

OR

- B**
- You have no problems walking about no problems with self care no problems with usual activities **moderate** pain or discomfort no anxiety or depression
 - You are **not** confident you can manage your condition
 - You can have a GP appointment in **one week's** time
 - You **rarely** see your friends or relatives

Please tick one box:

Choice A

☐

Choice B

☐

Question 23: If you had to choose between A and B below, which would you choose?

A

- You have **some** problems walking about
some problems with self care
no problems with usual activities
moderate pain or discomfort
moderate anxiety or depression
- You are **moderately** confident you can manage your condition
- You can have a GP appointment in **one week's** time
- You see friends or relatives **daily**

OR

B

- You have
no problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **not** confident you can manage your condition
- You can have a GP appointment in **three week's** time
- You see friends or relatives every **few days**

Please tick one box:

Choice A

☐

Choice B

☐

Question 24: If you had to choose between A and B below, which would you choose?

A

- You have
some problems walking about
some problems with self care
no problems with usual activities
moderate pain or discomfort
moderate anxiety or depression
- You are **moderately** confident you can manage your condition
- You can have a GP appointment in **three week's** time
- You **rarely** see your friends or relatives

OR

B

- You have
no problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
- You are **not** confident you can manage your condition
- You can have a GP appointment **tomorrow**
- You see friends or relatives **daily**

Please tick one box:

Choice A

☐

Choice B

☐

Question 25: If you had to choose between A and B below, which would you choose?

- A**
- You have **some** problems walking about **some** problems with self care no problems with usual activities **moderate** pain or discomfort **moderate** anxiety or depression
 - You are **not** confident you can manage your condition
 - You can have a GP appointment **tomorrow**
 - You see friends or relatives **daily**

OR

- B**
- You have no problems walking about no problems with self care no problems with usual activities **moderate** pain or discomfort no anxiety or depression
 - You are **totally** confident you can manage your condition
 - You can have a GP appointment in **one week's time**
 - You see friends or relatives every **few days**

Please tick one box:

Choice A ☐

Choice B ☐

Question 26: If you had to choose between A and B below, which would you choose?

- A**
- You have **some** problems walking about **some** problems with self care no problems with usual activities **moderate** pain or discomfort **moderate** anxiety or depression
 - You are **not** confident you can manage your condition
 - You can have a GP appointment in **one week's time**
 - You **rarely** see your friends or relatives

OR

- B**
- You have no problems walking about no problems with self care no problems with usual activities **moderate** pain or discomfort no anxiety or depression
 - You are **totally** confident you can manage your condition
 - You can have a GP appointment in **three week's time**
 - You see friends or relatives **daily**

Please tick one box:

Choice A ☐

Choice B ☐

MANY THANKS FOR TAKING THE TIME TO COMPLETE THIS FORM

Question 27: If you had to choose between A and B below, which would you choose?

- A**
- You have **some** problems walking about
some problems with self care
no problems with usual activities
moderate pain or discomfort
moderate anxiety or depression
 - You are **not** confident you can manage your condition
 - You can have a GP appointment in **three week's** time
 - You see friends or relatives every **few days**

OR

- B**
- You have no problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
 - You are **totally** confident you can manage your condition
 - You can have a GP appointment **tomorrow**
 - You **rarely** see your friends or relatives

Please tick one box:

Choice A ☐

Choice B ☐

Question 28: If you had to choose between A and B below, which would you choose?

- A**
- You have no problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
 - You are **moderately** confident you can manage your condition
 - You can have a GP appointment **tomorrow**
 - You **rarely** see your friends or relatives

OR

- B**
- You have **some** problems walking about
no problems with self care
no problems with usual activities
moderate pain or discomfort
no anxiety or depression
 - You are **not** confident you can manage your condition
 - You can have a GP appointment in **one week's** time
 - You see friends or relatives **daily**

Please tick one box:

Choice A ☐

Choice B ☐

MANY THANKS FOR TAKING THE TIME TO COMPLETE THIS FORM

Appendix G. Patient Questionnaires for Expert Patient Programme evaluation

Evaluation of the Expert Patient Programme

National Primary Care Research and Development Centre, University of Manchester, in conjunction with the Department of Health and the Expert Patients Programme



Dear Patient

You recently agreed to take part in our Evaluation of the Expert Patient Programme (EPP). We would be very grateful if you would complete the enclosed questionnaire, and return it to us in the prepaid envelope provided (no stamp is required).

Please answer ALL the questions.

If you have any problems completing the form, please call a member of the EPP evaluation team on 0161 275 7601 and we will be happy to assist.

The code number below will help us identify your questionnaire - there is no need to write your name or address on the form.

All information you provide will be treated in the strictest confidence

Many thanks for your help

Anne Kennedy

EPP evaluation team leader

Patient identification number

--	--	--	--

Please write in today's date below e.g. 1st May 2003 would be 01 05 03

Day:

Month:

Year:

Your Background

Are you:
(please tick one box)

Male ☐

Female ☐

Your date of birth:

(please write in the numbers below e.g. 1st May 1960 would be 01 05 60)

Day:

Month:

Year:

Who do you live with (if anybody) in your current home?
(Please tick all boxes that apply)

Live alone ☐

Spouse / partner ☐

Parent(s) ☐

Children under 18 ☐

Children over 18 ☐

Other family ☐

Non-family ☐

Is your accommodation:
(please tick one box only)

Owner-occupied / mortgaged? ☐

Rented from local authority / ☐
housing association?

Rented from a private landlord? ☐

Other arrangements? ☐
(please describe)

Is there a car or van normally available for use by you?

Yes ☐

No ☐

If yes, how many are normally available?

One ☐

Two or more ☐

Which of the following best describes you?
(Please tick one box only)

White - British ☐

Asian or Asian British ☐

Black or Black British ☐

Chinese or Chinese British ☐

Mixed ethnicity ☐

Other ☐

Which of these best describes your current work situation?
(Please tick one box only)

- In paid work
(including self-employed) ☐
- Unemployed ☐
- Retired from paid work ☐
- Unable to work because of
long-term disability or ill health ☐
- Looking after the family or home ☐
- In full-time education or training ☐
- Other ☐

Which of these qualifications do you have?
(Please tick all boxes that apply)

- 1 or more O levels / CSE / GCSEs (any grade) ☐
- 1 or more A levels or AS levels ☐
- Degree ☐
- NVQ ☐
- Other qualification (e.g. City and Guilds) ☐
- No qualifications ☐

Your long-term health problems

What would you say is your MAIN problem for which you are seeking help (the one that has the biggest effect on your day-to-day health)?

Diabetes ☐

Arthritis ☐

Heart Disease ☐

Asthma ☐

Other ☐
(Please describe in the box below)

How long have you had this condition? Years Months

What other conditions do you have?

Diabetes ☐

Arthritis ☐

Heart Disease ☐

Asthma ☐

Other ☐
(Please describe in the box below)

In which months are your long-term condition(s) at their worst?
(tick all that apply)

Much the same all year ☐

January to March ☐

April to June ☐

July to September ☐

October to December ☐

Your health and day-to-day activities

Your health

In general would you say your health is...
(please tick one box only):

- Excellent☐
- Very good☐
- Good☐
- Fair☐
- Poor☐

Your abilities

Please tick the one response for each question that best describes your usual abilities over the **past 4 weeks**:

Are you able to....
(please tick on box)

	Without any difficulty	With some difficulty	With much difficulty	Unable to do
Dress yourself, including tying shoelaces and doing buttons?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Brush/comb your hair?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Stand up from an armless chair?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Get in and out of bed?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Get up from off the floor?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Cut your food with a knife or fork?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Lift a full cup or glass to your lips?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Walk outdoors the length of a football field?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Walk outdoors for the length of several football fields?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Climb up five steps?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Are you able to....
(please tick one box)

	Without any difficulty	With some difficulty	With much difficulty	Unable to do
Climb up one flight of stairs?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Wash and dry your entire body?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Get on and off the toilet?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Take a bath?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Reach and get down a bag of sugar from just above your head?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Bend down (such as to pick up clothing from the floor)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Open jars which have previously been opened?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Turn taps on and off?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Run errands and shop?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Do household chores (such as vacuuming, gardening, laundry and DIY)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Get to places out of walking distance (by car or public transport)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Carry a bag of groceries across a room?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Your health and what you do

During the **past 4 weeks**, how much...
(please tick one box):

	Not at all	Slightly	Moderately	Quite a lot	Almost totally
Has your health interfered with your normal social activities with family, friends, neighbours, or groups?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Has your health interfered with your hobbies or recreational activities?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Has your health interfered with your household chores?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Has your health interfered with your errands and shopping?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Pain

Please circle the **one** number that best describes your physical discomfort or pain on the
average over **past 4 weeks**:

None

1234567891011121314151617181920

As bad as you
can imagine

Please circle the **one** number that best describes your physical discomfort or pain at its
worst over the **past 4 weeks**:

None

1234567891011121314151617181920

As bad as you
can imagine

During the past 4 weeks, how often have you had physical discomfort or pain?
(If you have had more than one discomfort or pain, answer by describing your feelings of discomfort or pain in general).
(please tick one box only)

- Never ☐
- Once or twice ☐
- A few times ☐
- Fairly often ☐
- Very often ☐
- Every day or almost every day ☐

How much bodily discomfort or pain have you generally had during the past 4 weeks? (please tick one box only)

- None ☐
- Very mild ☐
- Mild ☐
- Moderate ☐
- Severe ☐
- Very severe ☐

When you had physical discomfort or pain during the past 4 weeks, how long did it usually last? (If you have had more than one discomfort or pain, answer by describing your feelings of discomfort or pain in general).
(please tick one box only)

- Didn't have any ☐
- A few minutes ☐
- Several minutes to an hour ☐
- Several hours ☐
- A day or two ☐
- More than 2 days ☐

Your Energy

These questions are about how you feel and how things have been with you during the past month.

How much time during the past 4 weeks...
(please tick one box only)

	None of the time	A little of the time	Some of of the time	A good bit of the time	Most of the time	All of the time
Did you feel worn out?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Did you have a lot of energy?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Did you feel tired?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Did you have enough energy to do the things you wanted to?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Did you feel full of pep?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Your breathing

During the past 4 weeks, how much have you been troubled by shortness of breath when doing your normal daily activities?
(please tick one box only)

Not at all	<input type="checkbox"/>
Slightly	<input type="checkbox"/>
Moderately	<input type="checkbox"/>
Quire a bit	<input type="checkbox"/>
Almost totally	<input type="checkbox"/>

Your feelings about your health

These questions are about how you feel and how things have been with you during the past month.

How many times during the past 4 weeks....
(please tick one box only)

[illegible]

Your Exercise

During the past week (even if it was not a typical week), how much total time (for the entire week) did you spend on each of the following?
(please tick one box only)

Stretching or strengthening exercises (range of motion, using weights, etc.)

- None ☐
- Less than 30 minutes/week ☐
- 30-60 minutes/week ☐
- 1-3 hours/week ☐
- More than 3 hours/week ☐

Walk for exercise

- None ☐
- Less than 30 minutes/week ☐
- 30-60 minutes/week ☐
- 1-3 hours/week ☐
- More than 3 hours/week ☐

Swimming or aquatic exercise

- None ☐
- Less than 30 minutes/week ☐
- 30-60 minutes/week ☐
- 1-3 hours/week ☐
- More than 3 hours/week ☐

Bicycling (including stationary exercise bike)

- None ☐
- Less than 30 minutes/week ☐
- 30-60 minutes/week ☐
- 1-3 hours/week ☐
- More than 3 hours/week ☐

Other aerobic exercise equipment (Stairmaster, rowing or skiing machine)

- None ☐
- Less than 30 minutes/week ☐
- 30-60 minutes/week ☐
- 1-3 hours/week ☐
- More than 3 hours/week ☐

Other aerobic exercise – specify: _____

- None ☐
- Less than 30 minutes/week ☐
- 30-60 minutes/week ☐
- 1-3 hours/week ☐
- More than 3 hours/week ☐

Your smoking

Are you a current smoker?

- Yes ☐
- No ☐

If yes, how many cigarettes do you smoke on a typical day?

cigarettes per day

Your Health Today

By placing a tick in one box in each group below, please indicate which statement best describes your own health state today.
(Do not tick more than one box in each group)

Mobility

- I have no problems in walking about ☐
- I have some problems in walking about ☐
- I am confined to bed ☐

Self-Care

- I have no problems with self-care ☐
- I have some problems washing or dressing myself ☐
- I am unable to wash or dress myself ☐

Usual Activities

- I have no problems with performing my usual activities (e.g. work, study, housework, family or leisure activities) ☐
- I have some problems with performing my usual activities ☐
- I am unable to perform my usual activities ☐

Pain/Discomfort

- I have no pain or discomfort ☐
- I have moderate pain or discomfort ☐
- I have extreme pain or discomfort ☐

Anxiety/Depression

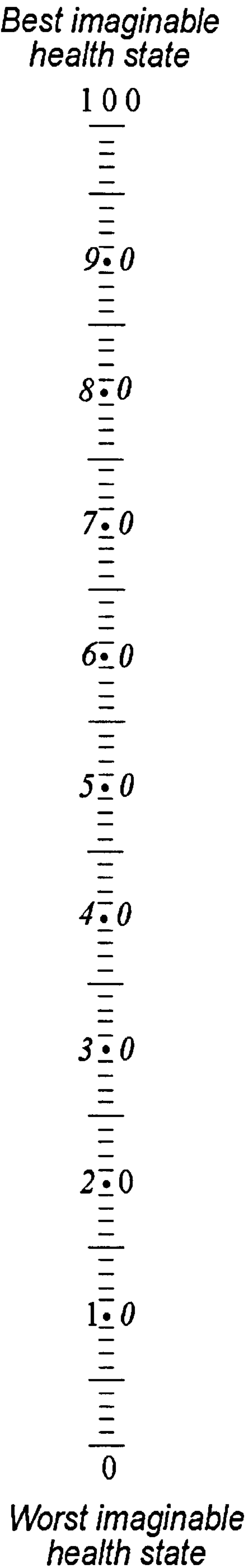
- I am not anxious or depressed ☐
- I am moderately anxious or depressed ☐
- I am extremely anxious or depressed ☐

Compared with my general level of health over the past 12 months, my health state today is:

- Better ☐
- Much the same ☐
- Worse ☐

- Please indicate on this scale how good or bad your own health is today.
 - The best health state you can imagine is marked 100 and the worst health state you can imagine is marked 0.
 - Please draw a line from the box below to the point on the scale that indicates how good or bad your health state is today.

Your own
health
state
today



Your confidence in doing activities

We would like to know **how confident** you are in doing certain activities.

For each of the following questions, please **circle** the number that corresponds to your confidence that you can do the tasks regularly at the present time.

How confident are you that you can.....

Having an illness often means doing different tasks and activities to manage your condition. How confident are you that you can do all the things necessary to manage your condition on a regular basis?

Not at all confident	1	2	3	4	5	6	7	8	9	10	Totally confident
-----------------------------	---	---	---	---	---	---	---	---	---	----	--------------------------

Judge when the changes in your illness mean you should visit a doctor?

Not at all confident	1	2	3	4	5	6	7	8	9	10	Totally confident
-----------------------------	---	---	---	---	---	---	---	---	---	----	--------------------------

Do the different tasks and activities needed to manage your health condition so as to reduce your need to see a doctor?

Not at all confident	1	2	3	4	5	6	7	8	9	10	Totally confident
-----------------------------	---	---	---	---	---	---	---	---	---	----	--------------------------

Reduce the emotional distress caused by your health condition so that it does not affect your everyday life?

Not at all confident	1	2	3	4	5	6	7	8	9	10	Totally confident
-----------------------------	---	---	---	---	---	---	---	---	---	----	--------------------------

Do things other than just taking medication to reduce how much your illness affects your everyday life?

Not at all confident	1	2	3	4	5	6	7	8	9	10	Totally confident
-----------------------------	---	---	---	---	---	---	---	---	---	----	--------------------------

How confident are you that you can.....

Reduce your physical discomfort or pain?

Not at all confident	1	2	3	4	5	6	7	8	9	10	Totally confident
---------------------------------	---	---	---	---	---	---	---	---	---	----	------------------------------

Keep the fatigue caused by your disease from interfering with the things you want to do?

Not at all confident	1	2	3	4	5	6	7	8	9	10	Totally confident
---------------------------------	---	---	---	---	---	---	---	---	---	----	------------------------------

Keep the physical discomfort or pain of your disease from interfering with the things you want to do?

Not at all confident	1	2	3	4	5	6	7	8	9	10	Totally confident
---------------------------------	---	---	---	---	---	---	---	---	---	----	------------------------------

Keep any other symptoms or health problems you have from interfering with the things you want to do?

Not at all confident	1	2	3	4	5	6	7	8	9	10	Totally confident
---------------------------------	---	---	---	---	---	---	---	---	---	----	------------------------------

Control any symptoms or health problems you have so that they don't interfere with the things you want to do?

Not at all confident	1	2	3	4	5	6	7	8	9	10	Totally confident
---------------------------------	---	---	---	---	---	---	---	---	---	----	------------------------------

How confident are you that you can.....

Do gentle exercises for muscle strength and flexibility three to four times per week (range of motion, using weights, etc)?

Not at all confident	1	2	3	4	5	6	7	8	9	10	Totally confident
---------------------------------	---	---	---	---	---	---	---	---	---	----	------------------------------

Do an aerobic exercise such as walking, swimming, or bicycling three to four times each week?

Not at all confident	1	2	3	4	5	6	7	8	9	10	Totally confident
---------------------------------	---	---	---	---	---	---	---	---	---	----	------------------------------

Exercise without making your symptoms worse?

Not at all confident	1	2	3	4	5	6	7	8	9	10	Totally confident
---------------------------------	---	---	---	---	---	---	---	---	---	----	------------------------------

How confident are you that you can.....

Keep from getting discouraged when nothing you do seems to make any difference?

Not at all confident	1	2	3	4	5	6	7	8	9	10	Totally confident
-----------------------------	---	---	---	---	---	---	---	---	---	----	--------------------------

Keep from feeling sad or down in the dumps?

Not at all confident	1	2	3	4	5	6	7	8	9	10	Totally confident
-----------------------------	---	---	---	---	---	---	---	---	---	----	--------------------------

Keep yourself from feeling lonely

Not at all confident	1	2	3	4	5	6	7	8	9	10	Totally confident
-----------------------------	---	---	---	---	---	---	---	---	---	----	--------------------------

Do something to make yourself feel better when you are feeling lonely?

Not at all confident	1	2	3	4	5	6	7	8	9	10	Totally confident
-----------------------------	---	---	---	---	---	---	---	---	---	----	--------------------------

Do something to make yourself feel better when you are feeling discouraged?

Not at all confident	1	2	3	4	5	6	7	8	9	10	Totally confident
-----------------------------	---	---	---	---	---	---	---	---	---	----	--------------------------

Do something to make yourself feel better when you feel sad or down in the dumps?

Not at all confident	1	2	3	4	5	6	7	8	9	10	Totally confident
-----------------------------	---	---	---	---	---	---	---	---	---	----	--------------------------

Your quality of life

This part of the questionnaire is designed to help us get a full picture of the quality of your life. We want you to tell us how satisfied or dissatisfied you are with a number of areas of your life.

Life in General

How do you feel about your life as a whole, today?
(Please tick one box only)

- Terrible ☐
- Displeased ☐
- Mostly dissatisfied ☐
- Mixed ☐
- Mostly satisfied ☐
- Pleased ☐
- Delighted ☐

Your Life Opportunities

In the past year, have there been times when you wanted to improve any of the following aspects of your life but were unable to?
(Please tick all boxes that apply)

- | | | | |
|-------------------|--------------------------|------------------|--------------------------|
| Work or education | <input type="checkbox"/> | Living situation | <input type="checkbox"/> |
| Finances | <input type="checkbox"/> | Family life | <input type="checkbox"/> |
| Leisure | <input type="checkbox"/> | Safety | <input type="checkbox"/> |
| Social life | <input type="checkbox"/> | Health | <input type="checkbox"/> |

Your work and education

If working: How do you feel about your job?
(Please tick one box only)

- Terrible ☐
- Displeased ☐
- Mostly dissatisfied ☐
- Mixed ☐
- Mostly satisfied ☐
- Pleased ☐
- Delighted ☐

If not working: How do you feel about not working?

(Please tick one box only)

- Terrible ☐
- Displeased ☐
- Mostly dissatisfied ☐
- Mixed ☐
- Mostly satisfied ☐
- Pleased ☐
- Delighted ☐

Your finances

How frequently (if at all) do you find it difficult to meet the cost of household bills?

(Please tick one box only)

- All of the time ☐
- Most of the time ☐
- Some of the time ☐
- Seldom ☐
- Never ☐

How do you feel about your financial situation?

(Please tick one box only)

- Terrible ☐
- Displeased ☐
- Mostly dissatisfied ☐
- Mixed ☐
- Mostly satisfied ☐
- Pleased ☐
- Delighted ☐

Your leisure

How do you feel about your leisure activities?

(Please tick one box only)

- Terrible ☐
- Displeased ☐
- Mostly dissatisfied ☐
- Mixed ☐
- Mostly satisfied ☐
- Pleased ☐
- Delighted ☐

Your social life

How do you feel about the number of friends you have?
(Please tick one box only)

- Terrible ☐
- Displeased ☐
- Mostly dissatisfied ☐
- Mixed ☐
- Mostly satisfied ☐
- Pleased ☐
- Delighted ☐

How do you feel about the quality of your friendship(s)?
(Please tick one box only)

- Terrible ☐
- Displeased ☐
- Mostly dissatisfied ☐
- Mixed ☐
- Mostly satisfied ☐
- Pleased ☐
- Delighted ☐

Your living situation

How do you feel about your accommodation?
(Please mark one box only)

- Terrible ☐
- Displeased ☐
- Mostly dissatisfied ☐
- Mixed ☐
- Mostly satisfied ☐
- Pleased ☐
- Delighted ☐

If living with other people: How do you feel about the people that you live with?
(Please mark one box only)

Terrible	<input type="checkbox"/>
Displeased	<input type="checkbox"/>
Mostly dissatisfied	<input type="checkbox"/>
Mixed	<input type="checkbox"/>
Mostly satisfied	<input type="checkbox"/>
Pleased	<input type="checkbox"/>
Delighted	<input type="checkbox"/>

If living alone: How do you feel about living alone?
(Please mark one box only)

Terrible	<input type="checkbox"/>
Displeased	<input type="checkbox"/>
Mostly dissatisfied	<input type="checkbox"/>
Mixed	<input type="checkbox"/>
Mostly satisfied	<input type="checkbox"/>
Pleased	<input type="checkbox"/>
Delighted	<input type="checkbox"/>

Your family

How often do you have contact with a relative (not including those who live with you) either face to face or by telephone?
(Please tick one box only)

Not at all	<input type="checkbox"/>	At least 3 monthly	<input type="checkbox"/>
Daily	<input type="checkbox"/>	At least yearly	<input type="checkbox"/>
At least weekly	<input type="checkbox"/>	Less than yearly	<input type="checkbox"/>
At least monthly	<input type="checkbox"/>		

How do you feel about your relationship with your family?
(Please tick one box only)

Terrible	<input type="checkbox"/>
Displeased	<input type="checkbox"/>
Mostly dissatisfied	<input type="checkbox"/>
Mixed	<input type="checkbox"/>
Mostly satisfied	<input type="checkbox"/>
Pleased	<input type="checkbox"/>
Delighted	<input type="checkbox"/>

Your safety

In the past year, have you been a victim of violence?

Yes ☐

No ☐

How do you feel about your personal safety?
(Please tick one box only)

- Terrible ☐
- Displeased ☐
- Mostly dissatisfied ☐
- Mixed ☐
- Mostly satisfied ☐
- Pleased ☐
- Delighted ☐

Your health

How do you feel about your health?
(Please tick one box only)

- Terrible ☐
- Displeased ☐
- Mostly dissatisfied ☐
- Mixed ☐
- Mostly satisfied ☐
- Pleased ☐
- Delighted ☐

How do you feel about your present mental health (nerves)?
(Please mark one box only)

- Terrible ☐
- Displeased ☐
- Mostly dissatisfied ☐
- Mixed ☐
- Mostly satisfied ☐
- Pleased ☐
- Delighted ☐

Life overall

How do you feel about your life as a whole?

(Please tick one box only)

Terrible	<input type="checkbox"/>
Displeased	<input type="checkbox"/>
Mostly dissatisfied	<input type="checkbox"/>
Mixed	<input type="checkbox"/>
Mostly satisfied	<input type="checkbox"/>
Pleased	<input type="checkbox"/>
Delighted	<input type="checkbox"/>

Your clubs and associations

Are you actively involved in any of the following clubs or associations?
(please tick all boxes that apply)

- Sports club ☐
- Neighbourhood watch scheme ☐
- Sports supporter club ☐
- Tenants' group ☐
- Social club ☐
- Residents' association ☐
- Volunteers e.g.
St Johns Ambulance ☐
- Local council ☐
- Hobby or interest group ☐
- Church or religious group ☐
- Political party ☐
- Other ☐
(give details)

.....

Your visits to the doctor

Thinking about the last visit you made to your doctor (GP or hospital doctor), do you agree or disagree with the following statements?
(please tick one box only)

	Strongly Agree	Agree	No opinion	Disagree	Strongly Disagree
I suggested a certain kind of treatment to my doctor	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I insisted on a particular kind of test or treatment for my symptoms	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I expressed doubt about the tests or treatment that my doctor recommended	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I gave my opinion on the pros and cons of the types of tests and treatment that my doctor ordered	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Your self-care activities

In which, if any, of the following ways do you currently help yourself to manage your condition, and how much do you do so?
(please tick one box only)

	Not at all	A fair amount	A great deal
Diet	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Alternative/ complementary products	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Complementary therapy	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Exercise	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Relaxation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Looking for information	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Do you belong to any patient support groups?

Yes ☐

No ☐

If YES, which ones?

.....

.....

.....

Your use of hospital services

Have you had any overnight stays in hospital during the last 6 months?

Yes ☐

No ☐

Have you attended a day hospital during the last 6 months?

Yes ☐

No ☐

Have you had any outpatient appointments during the last 6 months?

Yes ☐

No ☐

Have you attended an Accident & Emergency department during the last 6 months?

Yes ☐

No ☐

Have you attended hospital for day case surgery during the last 6 months?

Yes ☐

No ☐

Your use of services outside the hospital

Please estimate the total number of contacts for each of the service below, **during the last 6 months** (please enter '0' if a particular service was not used):

General Practitioner (at the surgery)

General Practitioner (at your home)

Practice Nurse (at GP surgery)

Occupational Therapist (at home)

District Nurse (at home)

Meals on Wheels

Physiotherapist (at home)

Other (please specify)

.....

Other (please specify)

.....

Your help from carers

Have you used any help (e.g. tasks around the home, shopping) from a carer (e.g. relative, friend) **in last 6 months?**

Yes ☐

No ☐

If **YES**, please estimate how many hours of care you have received **over the last 6 months**

--	--	--

 hours

Your use of day services

Have you used any day facilities (e.g. day centre, drop in centre, social club) during the last 6 months? Please include any private or voluntary services.

Yes ☐

No ☐

If YES, please state

- the type of service
- who provides the service
- the approximate number of times you attended
- how long you spent there each time you attended

TYPE OF SERVICE (e.g. day centre)	SERVICE-PROVIDING AGENCY (e.g. Health Authority, Social Services, Voluntary)	HOW MANY TIMES DID YOU ATTEND?	HOW LONG YOU WERE THERE EACH TIME ?
Drop in club	Voluntary organisation	3	60 mins

Your use of medication

Have you taken any prescribed any medication over the last 6 months?

Yes ☐

No ☐

If YES, please list the relevant medications administered during this period.

Name of medication	How many tablets do you take?	How often do you take the tablets?	What is the dose?	How long have you been taking the tablets?
Propanolol	2	Twice a day	5mg	6 months
Vertaxigen	1	When needed	50mg	3 weeks

Other costs you might have to pay

Have you had to pay for any prescription medication or over the counter drugs **during the last 6 months?**

Yes ☐

No ☐

If **YES**, approximately how much have you had to pay **during the last 6 months?**

£

--	--

Have you had to make any other payments as a result of your illness: e.g. improvements to your home?

Yes ☐

No ☐

If **YES**, approximately how much have you had to pay **during the last 6 months?**

£

--	--

Your work and your health

Are you in current employment?

Yes ☐

No ☐

If **YES**, please answer the following questions

If **NO**, please turn the page

How many hours do you work on average per week?

 hours

Have you had to take any days off from work over the last 6 months as a result of your illness?

Yes ☐

No ☐

If **YES**, please estimate the number of days of work you have missed as a result of your illness during this period

 days

Have you lost any earnings over the last 6 months as a result of your illness?

Yes ☐

No ☐

If **YES**, please estimate **the percentage** of your normal earnings you have lost during this period

 %

If you are not currently in paid employment, have you been employed for any period during the last 6 months?

Yes ☐

No ☐

If **YES**, please describe any previous employment in this period and the duration of employment during the period (e.g. 1 month, 2 weeks etc.)?

Please describe your first job

.....

Duration of employment (months or weeks)

.....

Please describe your second job if applicable

.....

Duration of employment (months or weeks)

.....

Have you given up any leisure time in the **last 6 months** as a result of your illness?

Yes ☐

No ☐

If **YES**, please estimate the number of hours lost during this period

hours

Your views about which group you would prefer to be in

When you return this questionnaire, we will randomise you to either go on an EPP course immediately or to wait for six months before going on a course.

We would like to know whether you have strong views about which group you would like to be in

Please circle the number on the scale below which best shows the group you would prefer to be in. If you don't mind when you start or are unsure please circle 5.

Start the course now	0	1	2	3	4	5	6	7	8	9	10	Wait 6 months to start course
---	---	---	---	---	---	---	---	---	---	---	----	--

For assistance, call the EPP evaluation team on 0161 275 7601

If you have any comments on the survey or any other aspects of the research, please write them here

If you have any problems completing the form, please call a member of the EPP evaluation team on 0161 275 7601 and we will be happy to assist.

MANY THANKS FOR YOUR HELP WITH THIS RESEARCH